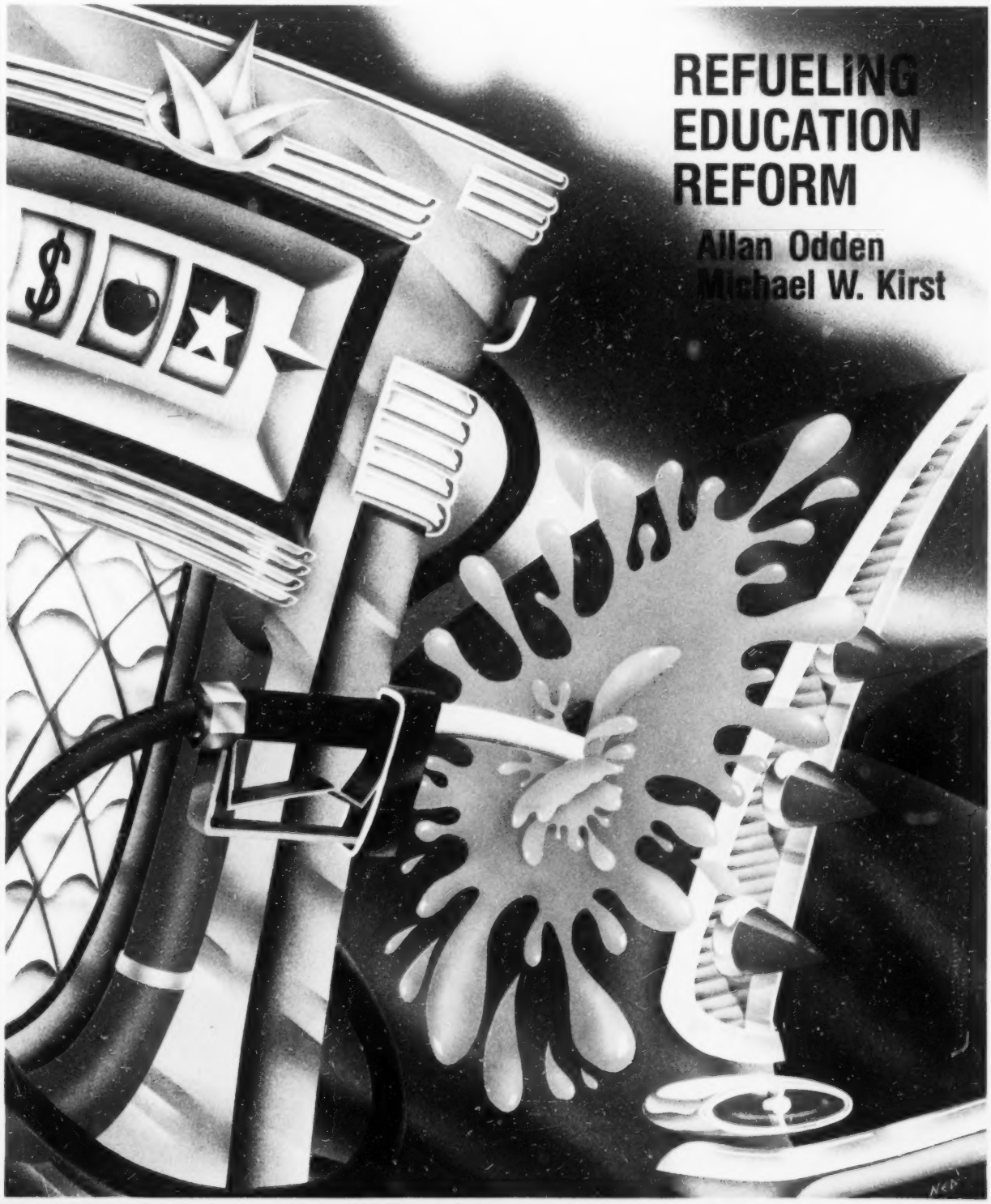
A8680-1A

PHI DELTA

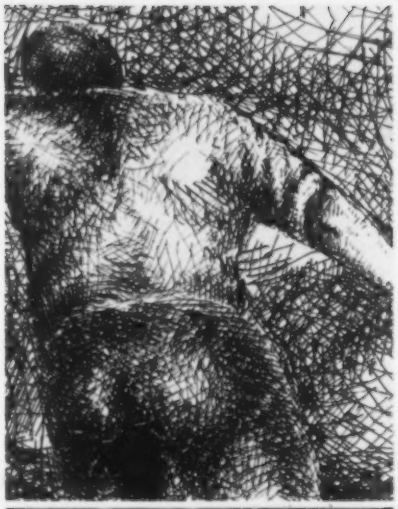
JANUARY 1986

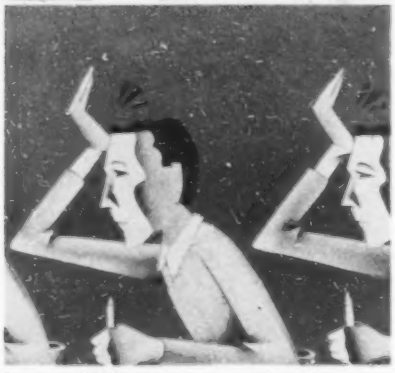
KAPPAN



REFUELING EDUCATION REFORM

Allan Odden
Michael W. Kirst





Phi Delta January 1986 • Volume 67 • Number 5

KAPPAN

FEATURES

335 **Sources of Funding for Education Reform,** *by Allan Odden*

341 **Sustaining the Momentum of State Education Reform: The Link Between Assessment and Financial Support,** *by Michael W. Kirst*

346 **Should a Million and a Half Students Be Required to Take the SAT Next Year?,** *by James Crouse*

353 **Perspectives for the Reform of Science Education,** *by Paul DeHart Hurd*

359 **The Role of Native-Language Instruction in Bilingual Education,** *by José A. Cárdenas*

364 **Illiteracy and Inner-City Unemployment,** *by Michael Bernick*

368 **Public and Nonpublic Schools: Finding Ways to Work Together,** *by Phyllis L. Blaunstein*

373 **The New Private Schools and Their Historic Purpose,** *by Patricia M. Lines*

380 **Learning Communities for Curriculum and Staff Development,** *by Anne Wescott Dodd and Eileen Rosenbaum*

384 **The Madison Workshops – One Approach to Staff Development,** *by Howard Henderson*

386 **A School/University Partnership That Fosters Inquiry-Oriented Staff Development,** *by Forrest W. Parkay*

390 **Of Hopes and Expectations,** *by Judy Meloy*

DEPARTMENTS

330 **THE EDITOR'S PAGE,** Welcome to Dolan

331 **WASHINGTON REPORT,** *by Anne C. Lewis*

333 **STATELINE,** *by Chris Pipho*

392 **RESEARCH,** *by Gerald W. Bracey*

396 **DE JURE,** Teachers' First Amendment Rights Eroding, *by Thomas J. Flygare*

398 **PROTOTYPES,** An Effective Reading Program in a Rural Setting, *by Gary A. Negin*

399 **BOOKS,** *Laurence D. Brown*

402 **NEWSNOTES**

405 **BACKTALK**

Cover Illustration: Ned Shaw

Sources of Funding for Education Reform

by Allan Odden

Improving the quality of education is a costly endeavor, says Mr. Odden. Recent appearances to the contrary, funds for reform are actually in short supply. Revenues for education will do well to stay even over the next five years.

LESS THAN THREE years have elapsed since the release of *A Nation at Risk* and the accompanying calls to improve U.S. public schools. Yet a number of state legislatures have already acted on the basic recommendations of that and other, similar reports. Indeed, the education reform movement has moved faster than any public policy reform in modern history.[1] All the states have expanded their school improvement programs, nearly all have increased high school graduation requirements, most have stiffened college admission requirements, many are deepening the content of course offerings, and many are enacting a variety of policies to strengthen the teaching profession.

Moreover, signs in many states indicate that the reforms are being implemented as intended. A number of studies in California, for example, have shown that students are attending school longer, taking more and tougher academic courses, receiving better counsel-

ALLAN ODDEN is an associate professor in the School of Education and director of the Southern California Policy Analysis for California Education (PACE) Center at the University of Southern California in Los Angeles.

If education reform has helped to reverse the declining finances of the public schools, it has not opened a gushing faucet.

ing about their secondary school programs and about college admission requirements, and scoring higher on achievement tests. Teachers are being paid more, and in some states they have new career-ladder options.[2] The swiftness of action, the breadth of the reform programs, and the multiple indicators of progress in the right direction give reason for optimism about the ultimate success of the education reform movement.

A key issue in maintaining the momentum of the reform movement, however, is funding. Improving the quality of education is a costly endeavor. Early studies estimated that revenues would need to increase by at least 20% in order to pay for most of the proposed reforms.[3] At the same time, attention to the simultaneous pursuit of fiscal equity and educational excellence was being urged.

Encouraging action has taken place in the struggle for fiscal equity. Of the 11 reform states that I discuss in this article, four – Arkansas, Georgia, Kentucky, and Texas – enacted fundamental changes in their school finance formulas as part of their education reform packages. Six of the other seven states had enacted school finance reforms during the 1970s; most of these reforms promoted the equalization of resources among school districts. So far, equity in school finance seems to have fared well in states that are actively involved in the reform movement.

Furthermore, funding levels for programs for students with special needs have also fared well in most reform states – sometimes receiving the same percent increase as the general aid formula, sometimes receiving a higher increase.[4] Although more analysis is needed on the ways in which funding changes have affected fiscal equity, the indicators above suggest that fiscal equity has not been forgotten in the pursuit of excellence – and indeed has shared center stage with excellence reforms in several states.

Less costly school improvement programs in the reform states have also expanded, both in number and in the amount of funding set aside for them. Missouri's recent reform program exemplifies the renewed attention being paid to initiatives designed to stimulate reform at the local level. Missouri enacted a set of education excellence initiatives to be financed at a level of $75 million from a separate fund. Programs supported by Missouri's Excellence in Education Fund include business/education partnerships, extended contracts for teachers and administrators, parent participation programs, instructional improvement projects, writing programs, high-technology projects, advanced placement programs, and opportunity classes for children at risk in grades 1 through 3. All are funded through some combination of state and local sources.

Finally, the gross indicators of funding for education reform look good. Arkansas, South Carolina, and Tennessee increased the sales tax by a penny in order to finance their reform programs. California and Texas each added more than a billion dollars to education in the first year of the reform, and Florida, Georgia, Missouri, and New York increased state aid to education by larger amounts than at any time in history. Kentucky's reform bill costs more than $300 million – just less than 30% of current state aid.

Now that the U.S. has several years of experience with the new wave of education reform, however, a more detailed look at the course of school funding suggests that optimism for fiscal growth needs to be restrained. Funds for education reform are actually in short supply. Two recent studies concluded that revenues for education – even in many reform states – will do well to stay even over the next five years, after adjustments for enrollment increases and inflation.[5] In the remainder of this article I will investigate the level of revenues that various sources of school financing have produced in recent years, and I will conclude with a prognosis for the fiscal needs of public education for the rest of the Eighties.

FUNDING PATTERNS

Nationwide, school funding has reversed the declines of the early 1980s and has begun to increase moderately.[6] Expenditures per pupil have increased 15.8% in the two years since the education reform movement started in 1983. When adjusted for inflation, the increase was only 7.2% – not the 20% that most of the reform plans will cost, but still better than the real losses of the preceding three years. Federal revenues have stayed about the same in nominal terms throughout the Eighties but have dropped in real terms from $6.7 billion in 1982 to $5.8 billion in 1985 (in 1979 inflation-adjusted dollars). During this period state and local revenues used for education increased in both real and nominal terms. Adjusted for inflation, state funds rose by 5.4%, and local funds rose by 6.4%. While there is comfort in the trends these figures re-

"I'm sorry I lost my diploma, but I can show you my father's cancelled checks."

veal, the funding increases are modest. If education reform has helped to reverse the declining finances of the public schools, it has not opened a gushing faucet. The figures for per-pupil expenditures would be even lower had enrollments been rising during the early 1980s – as they are now beginning to do.

The individual states that have enacted comprehensive reform programs appear, at first blush, to have improved funding for education dramatically above the national averages. Tables 1 and 2 depict levels of funding and changes in funding for 11 education reform states between 1982-83 and 1985-86. Several of the changes are impressive. Arkansas increased state aid by 18.7% in the first year of its reforms, by 18.3% in the following year, and by 10.4% in the third year. In the three years since its reform program began, California increased aid by $3.2 billion – a considerable amount by any standard. New York increased state aid by $613 million between 1983-84 and 1985-86, the largest aid increase in that state's history. When it enacted its reforms, South Carolina hiked state aid by 32.4%, bumping the total state and local revenues by 22.5%. Tennessee also increased state revenues by more than 20%, and Texas combined an increase in state aid of nearly $1 billion with a local revenue increase of nearly $500 million to bring total state and local increases to 19.4%. By any reckoning, these funding changes are large.

But if we analyze these funding changes over the long term, identify the trade-offs between state and local revenue changes, and adjust for enrollment growth and inflation, we find them modest overall. Consider the following examples. When California's funding increases are adjusted for enrollment increases (now averaging 100,000 students per year) and for inflation, revenues per pupil stayed even between 1984-85 and 1985-86, the third year of reform. Inflation and enrollment growth in California will require an additional $1.5 million dollars above what the reforms produced for each of the next few years in order to maintain a steady fiscal state. In South Carolina and Texas, large one-year increases were preceded and followed by lean years; in fact, with rising enrollments in Texas, state aid per pupil in nominal terms will drop in 1985-86 and in 1986-87. The state aid increase in New York between 1983-84 and 1984-85 – the largest in the state's history – seems to be almost completely offset by a lack of increase in local revenues for that year; in fact, combined state and local revenues increased by a larger amount in the year preceding that historic rise in state revenues. Put differently, New York's educational system fared better when state aid increases were more modest. Missouri's recent large increase in state aid came after three years of very small increases, though increased education revenues from the sales tax initiative (Proposition C) boosted education funding in the two years preceding the 1985 reforms.

The analysis above is not meant to criticize the Herculean efforts to increase education funding in the reform states. Rather, my point is that, even with extra effort, school funding has not increased all that much. Enrollment increases – a new phenomenon after a

Table 1.
Level of State and Local Revenues for Public Schools, 1982-83 to 1985-86

	1982-83	1983-84	1984-85	1985-86 (est.)
	(millions of dollars)			
Arkansas				
State	509.8	605.3	716.3	790.7
Local	272.2	293.4	301.4	331.4
Total	782.0	898.7	1,017.7	1,122.1
California				
State	8,052.6	9,328.4	10,443.6	11,227.3
Local	2,675.0	2,886.0	3,289.1	3,459.9
Total	10,727.6	12,214.4	13,732.7	14,687.2
Florida				
State	1,898.9	2,122.9	2,353.0	2,575.4
Local	1,140.8	1,274.4	1,422.0	1,532.1
Total	3,039.7	3,397.3	3,775.0	4,107.5
Georgia				
State	1,367.0	1,460.0	1,630.0	1,755.0
Local	n.a.	n.a.	n.a.	n.a.
Total	n.a.	n.a.	n.a.	n.a.
Illinois				
State	2,103.2	2,236.1	2,352.9	2,697.9
Local	2,974.4	3,182.9	3,208.0	n.a.
Total	5,077.6	5,419.0	5,560.9	n.a.
Kentucky				
State	1,048.0	1,141.8	1,160.5	1,179.2
Local	279.5	314.8	359.0	409.3
Total	1,327.5	1,456.6	1,519.5	1,588.5
Missouri				
State	827.7	976.0	1,019.2	1,165.7
Local	1,058.9	1,228.2	n.a.	n.a.
Total	1,886.6	2,204.2	n.a.	n.a.
New York				
State	4,643.9	4,877.0	5,490.0	5,893.0
Local	6,459.7	7,085.0	7,130.0	7,531.9
Total	11,103.6	11,962.0	12,620.0	13,424.9
South Carolina				
State	745.9	803.8	1,064.1	1,117.3
Local	453.8	519.9	557.8	585.7
Total	1,199.7	1,323.7	1,621.9	1,703.0
Tennessee				
State	774.4	773.2	938.2	1,004.0
Local	706.8	741.7	778.8	817.7
Total	1,481.2	1,514.9	1,717.0	1,821.7
Texas				
State	3,620.0	3,734.6	4,701.8	4,851.2
Local	3,349.0	3,673.1	4,142.4	n.a.
Total	6,969.0	7,407.7	8,844.2	n.a.

Figures are not comparable across states. State revenues in Missouri include Proposition C funds. California figures exclude lottery revenues.

decade and a half of enrollment declines – will require a great many new dollars. Inflation, though modest, requires extra funding. State aid increases often can supplant local funds so that the net funding increase is modest. And large hikes in one year may not set a pattern for subsequent years. Over time, even in education reform states that have experienced the largest funding increases, these factors produce changes that are modest, that barely keep funding for the educational system from losing ground. More detailed analyses of the sources of these recent increases in funding for schools provide insight into the reasons why funding for education is unlikely to rise dramatically in the near future – even in reform states.

SOURCES OF FUNDS FOR REFORM

Federal revenues. Revenues for public schools come from federal, state, and local governments, as well as from individuals and organizations in local communities. As I mentioned above, over the past few years federal education revenues have increased slowly in nominal terms and actually dropped when adjusted for inflation. Each year Congress introduces new education bills that would increase federal funding for education, but the bills usually do not make it to floor debate. Unless some dramatic change occurs in the issues pressing Congress, large federal deficits and growing noneducation expenditures make significant increases in federal aid to education an unlikely source of major new education revenues over the next five years.

State revenues. As Tables 1 and 2 show, the states have been stalwart providers of recent increases in funding for education, and this pattern holds for the last decade and a half. States became the primary sources of education funding in part because of the school finance reforms of the 1970s. States also contributed three of every five new dollars for schools during the recession of the early 1980s.[7] Furthermore, states have been the fiscal engine that has powered the education reforms enacted since 1983.

Indeed, changes in state taxes proved to be the major source of funds for expensive education reforms – both directly, through such vehicles as sales tax increases dedicated to school reform (as in Tennessee, Arkansas, and South Carolina), and indirectly, through tax increases to balance state budgets.[8] Because of these actions, states claimed to be leaders in financing key domestic programs, including education. These politically courageous actions, coming at a time when the fervor to cut taxes and expenditures had not yet clearly waned, helped set the stage for a robust response to the funding needs of the proposed reforms.

The states that have enacted comprehensive education reforms usually raised state taxes to finance them. Five states increased the sales tax and devoted nearly all the extra revenues to education. In Arkansas, the December 1983 reform package increased the sales tax by a penny; the yield from that tax hike – $67 million for the last six months of the 1984 fiscal year – helped cover the $95.2 million increase in state aid in the first year of the reforms. In the next year, however, the extra penny brought in more than $155 million,

Table 2.
Changes in State and Local Revenues for Public Schools, 1982-83 to 1985-86

	1982-83 to 1983-84 [millions of dollars (% change)]	1983-84 to 1984-85 [millions of dollars (% change)]	1984-85 to 1985-86 [millions of dollars (% change)]
Arkansas			
State	95.2 (18.7)	111.0 (18.3)	74.4 (10.4)
Local	21.2 (7.8)	8.0 (2.7)	30.3 (18.0)
Total	116.7 (14.9)	119.0 (13.2)	104.4 (10.3)
California			
State	1,275.8 (15.8)	1,152.0 (12.0)	783.7 (7.5)
Local	211.0 (7.9)	403.1 (14.0)	170.8 (5.2)
Total	1,486.8 (13.9)	1,518.3 (12.4)	954.5 (7.0)
Florida			
State	224.0 (11.8)	230.1 (10.8)	222.4 (9.5)
Local	133.6 (11.7)	147.6 (11.6)	110.1 (7.7)
Total	357.6 (11.8)	377.7 (11.1)	332.5 (8.8)
Georgia			
State	43.0 (6.8)	170.0 (11.6)	125.0 (7.7)
Local	n.a.	n.a.	n.a.
Total	n.a.	n.a.	n.a.
Illinois			
State	132.9 (6.3)	116.8 (5.2)	345.0 (14.7)
Local	208.5 (7.0)	25.1 (0.8)	n.a.
Total	341.4 (6.7)	141.9 (2.6)	n.a.
Kentucky			
State	93.8 (9.0)	18.7 (1.6)	18.7 (1.6)
Local	35.3 (12.6)	44.2 (14.0)	50.3 (14.0)
Total	129.1 (9.7)	62.9 (4.3)	69.0 (4.5)
Missouri			
State	148.3 (17.9)	43.2 (4.4)	146.5 (14.4)
Local	169.3 (16.0)	n.a.	n.a.
Total	317.6 (16.8)	n.a.	n.a.
New York			
State	233.1 (5.0)	613.0 (12.6)	403.0 (7.3)
Local	625.3 (9.7)	45.0 (0.6)	401.9 (5.6)
Total	858.4 (7.7)	658.0 (5.5)	804.9 (6.4)
South Carolina			
State	57.9 (7.8)	260.3 (32.4)	53.2 (5.0)
Local	66.1 (14.6)	37.9 (7.3)	27.9 (5.0)
Total	124.0 (10.3)	298.2 (22.6)	81.1 (5.0)
Tennessee			
State	-1.2 (-0.2)	165.0 (21.3)	65.8 (7.0)
Local	34.9 (4.9)	37.1 (5.0)	38.9 (5.0)
Total	33.7 (2.3)	202.1 (13.3)	104.7 (6.1)
Texas			
State	114.6 (3.2)	967.2 (25.9)	149.4 (3.2)
Local	324.1 (9.7)	469.3 (12.8)	n.a.
Total	438.7 (6.3)	1,436.5 (19.4)	n.a.

while state aid for elementary and secondary education went up $110 million. South Carolina gave education the entire one-cent increase in its sales tax, estimated to yield $202 million in 1985, $227 million in 1986, and $242 million in 1987. Tennessee's one-cent sales tax increase produced about $325 million in 1985, with elementary and secondary education receiving a hike of $165 million.

To garner public support for a one-cent tax increase in 1982, Missouri promised to use half of the proceeds to roll back local property tax levies. The rest, though, provided healthy revenue increases for education, boosting state aid by 18% in 1984. However, increases in education aid from general state revenues stagnated in the two years after Proposition C. Thus the hike in the sales tax supplanted a portion of natural increases in state aid.

Finally, Texas increased the sales tax modestly (from 4% to 4.125%) and expanded its base to provide the largest new, individual revenue source ($305 million) for its multi-billion-dollar education reform. During a special session of the legislature in 1984, Texas lawmakers increased several other taxes as well (including those on motor fuels, corporate franchises, insurance companies, and on the sale or rental of motor vehicles) to raise an extra $4.8 billion over three years to improve the state's schools and highways. From these sources, $2.7 billion was allocated for education, $1.4 billion for highways, and $700 million to balance the budget.

Illinois and Kentucky also increased state tax rates to fund education reform, but on a series of smaller items. Illinois raised a number of excise taxes (a 5% tax on long-distance telephone calls and an eight-cent cigarette tax) to fund its $92.5 million reform package. Kentucky restructured the depreciation schedules for its corporate income tax and increased the business inventory and corporation license taxes to help fund its $307 million reform program.

California and Florida enacted significant modifications in various state taxes to raise the revenues needed to finance the reforms the states enacted in the summer of 1983. By making state tax codes conform to federal codes and by putting property on the tax roll when it is completed or using the new market value immediately when property changes hands, California raised an estimated $400 million that, when added to the $350 million that both houses of the state legislature had already agreed to add to education, nearly covered the $800 million first-year cost of the reforms. Florida made numerous adjustments to the way business income is taxed and to the general sales tax, in order to produce the extra $100 million it needed for education reform.

Georgia and Missouri – the two states that enacted no changes in the tax code and funded their reforms from natural revenue growth – actually delayed implementation of the reforms for at least one fiscal year. Time will tell whether available revenues a year hence will prove adequate.

Finally, three states seem to be banking heavily on lottery revenues to give education a fiscal boost. Missouri hopes that its new lottery, to begin in 1986, will produce funds to finance future costs of reform. California enacted a lottery by initiative in the fall of 1984; it began operating in October 1985. Net proceeds are expected to be approximately $400 million ($243 million for the current school year), which is about what is needed to "cover" one year of student enrollment increases – hardly a major revenue boost. Finally, the large increases in state aid in New York occurred during the year in which the lottery produced significant new revenues; but the next year, when actual lottery revenues exceeded estimates by $43 million, the general state appropriation for education was reduced by that exact amount. Furthermore, as noted above, when state aid increased in 1985, local revenues fell, and the net gain was minimal. So in the first year, lottery revenues supplanted local funds, and in the next year, lottery funds supplanted state funds.

In short, education reform states have

"What money-making enterprise are you majoring in?"

The states have been stalwart providers of recent increases in funding, and this pattern holds for the last decade and a half.

been quite aggressive in finding new sources of revenue for public schools. They have increased tax rates, expanded tax bases, adopted lotteries, and pledged natural revenue increases for the schools. They have taken these steps at a time when tax increases are even less popular than usual, though public support seems to be available if tax increases are specifically aimed at improving the quality of public education. In the main, though, the additional revenues have given education a one- to two-year fiscal shot in the arm, after which revenues per pupil will do well to stay even in real terms. In short, while the states have creatively tapped a variety of sources to fund education reform, they have not opened wide the fiscal floodgates. At a time when government growth is limited, it seems that state dollar increases to reform and improve the education system have so far been modest.

Local revenues. As I mentioned above, local revenues have also played an important role in education finance in recent years. After the proliferation of property tax relief policies in the late 1970s, states turned once again – in the depths of the recession of the early 1980s – to the local property tax to help finance schools. Indeed, the large property tax increases between 1980 and 1982 were viewed with alarm by those who saw them as a reversal of gains made in the previous decade. Although the property tax continues to play a formidable role in school funding, it is unlikely to provide dramatic increases through higher tax rates, except perhaps in places where the tax is significantly underused. None of the education reforms discussed above, for example, was accompanied by mandatory increases in the property tax rates of the same order of magnitude as increases in state taxes, which changed as much as

25% in some states. In Arkansas, Florida, and Texas, however, modest property tax increases were required, and in Missouri requirements for increasing local property tax rates were loosened. Nevertheless, the property tax is still an unpopular tax,[9] and efforts to increase its use, except in times of severe fiscal stress, are likely to produce a backlash.[10] Thus higher property taxes are unlikely to be a source of new revenues for the schools.

There are several other potential sources of local revenue growth, however. The first includes local option sales and income taxes. Although these would create technical, administrative, and legal problems, they nevertheless are productive avenues to pursue for increasing local education revenues. In Tennessee, for example, the sales tax provides about 40% of local school revenues. Indeed, the popularity of the sales tax at the state level usually carries over to the local level as well; when put to a vote, taxpayers often support a local sales tax. Although local sales and income taxes are not widespread today, a growing number of communities are opting for their use, and they offer the potential for generating significant revenue in such states as California and Washington, where it is virtually impossible to increase local property taxes.

Other sources of local revenues include nontax income derived from various types of entrepreneurial activities. Lionel Meno identified three major sources of revenue of this kind:[11]

- donor funds, including direct cash donations to local districts, indirect cash donations through locally created educational foundations, and donations of goods and services;
- enterprise activities, such as leasing school services and facilities or charging user fees for various school materials and activities; and
- shared or cooperative activities with community colleges, other colleges and universities, and local government agencies, including sharing buses, parks, recreational centers, and pools.

Meno conducted an extensive study of the extent, nature, and amount of revenue produced by each of these categories of activities. He concluded that a district that used every one of them could possibly increase its budget by 9%. Meno found, however, that actual dollars raised varied from $1.57 per pupil to $75.57 per pupil. The largest actual revenue yield was about an additional 2% of the nationwide average expenditure per pupil.

In another study, Richard Yong and Alan Hickrod reached similar conclusions about the revenue-raising potential of entrepreneurial fund-raising activities.[12] Although the sums actually raised are not insignificant, the fact is that they do not offer much potential for enhancing local district revenues. Furthermore, many local enterprise activities entail new, shared school governance mechanisms – an understandable requirement, but perhaps too high a price to pay for the revenues that they yield.

On the other hand, the impact of well-organized local district educational foundations needs to be measured in terms beyond the revenues they yield directly. Foundation funds can expand community involvement in the schools, raise the interest of local businesses, and strengthen school/business partnerships. Noneducators involved in such foundations often develop a renewed appreciation for local schools and realize that schools do a pretty good job, given the complexity of the tasks they face and the level of resources they have available. Such activities can rekindle local support for schools that, over time, can become support for increased local funding and a catalyst for improving the schools. Viewed in this light, the seeming rise of local educational foundations across the U.S. could have significant positive effects on school funding.

NO MAGIC or secret sources of new funds for the schools exist. Local property taxes, state revenues from a number of tax sources, and federal aid will continue to provide the bulk of financing for our schools. These revenue sources have halted their decline of three years ago and are now increasing at modest rates. If the strength of education reform has reversed fiscal decline, though, it has not been sufficient – at least, given the fiscal evidence so far – to place funding for education on a sound course of improvement.

These current conditions, coupled with demographic trends for the future, make raising the level of education funding a challenge. Student enrollments are now on the rise; thus schools will need more money just to stay even with expenditures per pupil. More classrooms will be needed to house these students – another costly outlay. The rising public school enrollments will include larger numbers and percentages of minority, limited-English-proficient, poor, and learning-disabled students.[13] All these special categories of students will require extra services to meet their educational needs. And the quality of education for the rest of the population, though better than it was a few years ago, still needs to improve.

We must continue to increase school funding. Recent evidence shows that the educational system can change quickly when public demands for better quality are stated clearly and money for reform is provided. Continuing to improve the quality of public education will require sustained funding; providing education for a growing number of ever more diverse students will require sustained funding. If the nation responds in the next five years as it has in the past three, school funding can increase, and the quality of education can improve. We really have no other options.

---

1. Susan Fuhrman and Lorraine McDonnell, "The Political Context of School Reform," in Van Mueller and Mary McKeown, eds., *State Reform of Elementary and Secondary Education* (Cambridge, Mass.: Ballinger, forthcoming).
2. James Guthrie and Michael Kirst, eds., *Conditions of Education in California, 1985* (Berkeley: University of California, Policy Analysis for California Education, 1985).
3. Allan Odden, "Financing Educational Excellence," *Phi Delta Kappan*, January 1984, pp. 311-18.
4. Allan Odden, *Education Finance in the States: 1984* (Denver: Education Commission of the States, 1984).
5. Allan Odden, "Education Finance 1985: Rising Tide or Steady Fiscal State?," *Educational Evaluation and Policy Analysis*, in press; and Jack Ossman, *Projections of Education Expenditures and Revenues in California to 1990* (Berkeley: University of California, Policy Analysis for California Education, 1985).
6. Odden, "Education Finance 1985. . . ."
7. Odden, *Education Finance in the States: 1984*, p. 2.
8. Steven Gold and Corina L. Eckl, *State Fiscal Conditions Entering 1984* (Denver: National Conference of State Legislatures, Legislative Finance Paper No. 42, 1984).
9. Advisory Commission on Intergovernmental Relations, *Public Attitudes on Governments and Taxes: 1984* (Washington, D.C.: ACIR, 1984).
10. Steven Gold, "State Tax Increases of 1983: Prelude to Another Tax Revolt," *National Tax Journal*, March 1984, pp. 9-22.
11. Lionel R. Meno, "Sources of Alternative Revenues," in L. Dean Webb and Van D. Mueller, eds., *Managing Limited Revenues* (Cambridge, Mass.: Ballinger, 1984).
12. Richard Yong and G. Alan Hickrod, *Private Sector Support of K-12 Education: A Review of Selected Programs in Seventeen States and Recommendations for Illinois* (Normal: Illinois State University, Center for the Study of Educational Finance, April 1985).
13. Harold Hodgkinson, *All One System* (Washington, D.C.: Institute for Educational Leadership, 1985). K

Should a Million and a Half Students Be Required to Take the SAT Next Year?

by James Crouse

Most colleges could ignore their applicants' scores on the Scholastic Aptitude Test without appreciably altering the overall accuracy of their admissions decisions. The SAT is statistically redundant, Mr. Crouse claims.

BY THE TIME they graduate, nearly 1.5 million members of the high school class of 1985-86 will have taken the Scholastic Aptitude Test (SAT). Many of them will have taken the test more than once. All but a handful will take it because they wish to attend college and need to submit their scores to one of the institutions that require SAT scores for admission. Will these students have taken the test for any useful purpose?

No one knows the full costs of the SAT in both economic and human terms, but they must be enormous. The entire SAT enterprise would seem justifiable only if the use of the test resulted in some substantial benefits. The College Board and the Educational Testing Service (ETS) have claimed for decades that the SAT helps colleges make better academic decisions about which students to admit.

JAMES CROUSE is a professor of educational studies and sociology at the University of Delaware, Newark. *This article is adapted from his article in the May 1985 issue of the* Harvard Educational Review, *titled "Does the SAT Help Colleges Make Better Selection Decisions?" This paper was first presented at the 1984 Educational Testing Service seminar, Princeton, N.J.*[1]

In the May 1985 issue of the *Harvard Educational Review*,[2] I argued that the view of the benefits of the SAT held by the College Board and ETS is not supported by the facts. My analysis raises serious questions about how colleges can continue to justify requiring their applicants to take the test, about how ETS and the College Board can continue to justify promoting its use, and about how the American public can get better information about the test's purposes and whether they are met.

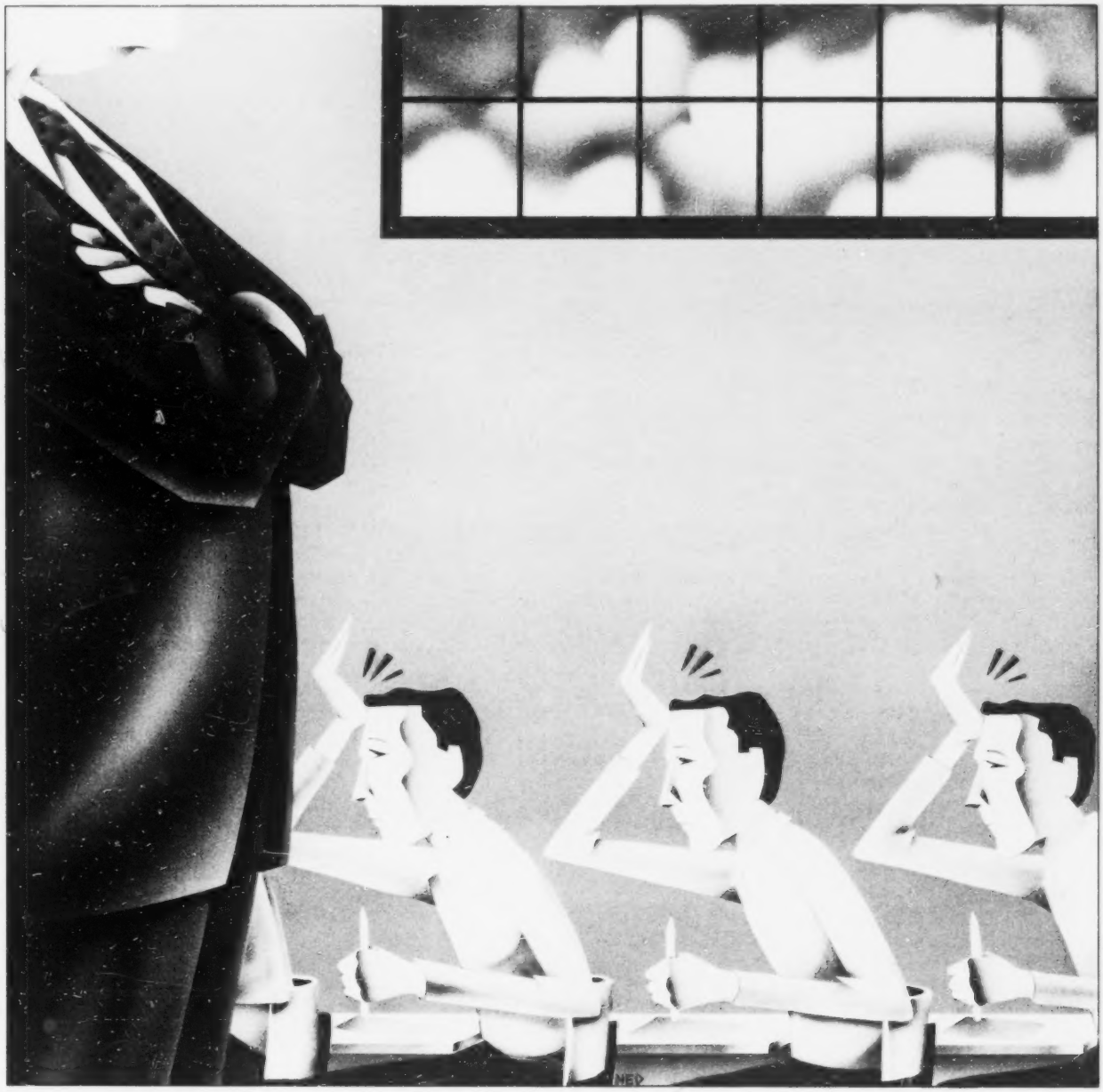
PURPOSES OF THE SAT

The College Board says that SAT scores are "intended to supplement the secondary school record and other relevant information about the student in assessing competence for college work."[3] Why use SAT scores as a supplement to assess competence for college work? Because, says ETS, though "students' previous grades are the most important indicator . . . the addition of scores on good tests adds significantly to the prediction of success."[4] ETS offers a significant amount of evidence to support its case.

The predictive effectiveness of the SAT has been assessed by the College Board, by ETS, and by others almost from the beginning of its use. In addition, since 1964 the College Board's Validity Study Service has encouraged colleges to conduct their own studies of the predictive effectiveness of the SAT, and slightly more than half of the four-year colleges and universities contacted in a 1979 survey by the College Board and the American Association of Collegiate Registrars and Admissions Officers had conducted some kind of study.[5] The College Board periodically releases results from these studies, which now number more than 3,000.[6]

Other organizations conducted nationally representative surveys of high school students in 1960, 1972, and 1980 to assess the state of U.S. education, and SAT scores are in these data files or can be added to them. Because each of these surveys follows students after high school, the predictive effectiveness of the SAT for college and later life can be measured.[7]

The results of these studies show overwhelmingly that the high school record and the SAT taken together are more effective in predicting freshman grades than the high school record alone. Data from the National Longitudinal Study (NLS) of the high school



class of 1972 indicate that students with the same high school class rank whose SAT scores differ had widely differing freshman academic records (see Table 1). A student with a higher SAT score had a higher predicted freshman grade-point average (GPA), freshman percentile rank in class, and probability of graduation. That student also had a smaller chance of earning average grades lower than B or C.

At first, these results struck me as dramatic evidence of the predictive effectiveness of the SAT. They seem to affirm the claims of the College Board and ETS and to refute critics who argue that the SAT adds little to predictions of college success based solely on the high school record.[8] But in fact these results, with which ETS and the College Board have justified colleges' use of the SAT for nearly 60 years, tell us almost nothing about the usefulness of the test to colleges.

Why? The simple reason is that prediction of freshman grades is not the ultimate goal of the admissions process. The purpose of college admissions is to make admissions decisions. The SAT can be useful to college admissions officers only if it helps them admit applicants whom they would reject using only high school records and helps them reject applicants whom they would admit using only high school records. If colleges admitted and rejected the same students, whether or not they used the SAT, the test would be redundant and could add nothing to actual college admissions decisions. This is in fact nearly the case.

REDUNDANCY OF THE SAT

We must assume that, if colleges predicted their applicants' freshman grades (for purposes of admission) using high school rank alone, some pre-

Illustration by Ned Shaw

When the yardstick used for success is college completion, the predictive value of the SAT is even less impressive.

dictions would be different from those made using high school rank together with SAT scores. How different would these predictions be? Would the difference be large enough to influence any college's actual admissions decisions?

To find out, let us look at the distribution of predicted freshman grades using high school rank alone to see how much it differs from the distribution of predicted freshman grades using both high school rank and SAT scores. I analyzed a sample of 2,781 high school seniors in the NLS survey who applied to four-year colleges as their first choice and for whom SAT scores, high school rank, and selectivity data for their first-choice colleges were available. I then predicted freshman grades for these applicants at their first-choice colleges in two ways. First, I used high school class rank alone; second, I used a combination of high school class rank and SAT scores. Table 2 presents these two distributions of predicted freshman GPAs and their cross-tabulation.

Table 2 is a composite of predicted grades for applicants at many different colleges. The reader should not fall into the trap of thinking that the distributions and their cross-tabulation in Table 2 are therefore unlike those of applicants to any individual college. In fact, I have examined versions of Table 2 for many different individual colleges, and these versions never differ in any important ways from Table 2. Consequently, I combined all colleges in the NLS to describe the applicant pool at the typical American college.

The distributions of these predicted freshman grades have some important consequences for understanding the usefulness of the SAT in making decisions about college admissions. Suppose that the typical college in the NLS decided to accept applicants whose predicted freshman GPA was above 2.5 and to reject those whose predicted freshman GPA was 2.5 or below. When we compare the effects of basing these predictions either on high school record alone or on high school rank plus SAT score, we can see that 74.2% of the applicants would be accepted and 16.6% rejected under either policy. Identical admissions decisions would be made on a total of 90.8% of the applicants. Therefore, only 9.2% of applicants to college would be accepted using one plan and rejected using the other.

Note that this finding does not depend on where the college sets the cutoff point for admissions. Setting the cutoff point at 2.0 or 3.0 will change the percentages somewhat but will still demonstrate that there is little to choose between a high-school-rank policy and a high-school-rank-plus-SAT policy. This finding is virtually guaranteed by the fact that the two distributions of predicted freshman grades correlate .881.

Thus colleges make identical admissions decisions, either to admit or reject, on more than 90% of their applicants whether they use high school grades alone or high school grades plus the SAT. The SAT, therefore, has little impact on who is in colleges' freshman classes and who is not. It is, in effect, statistically redundant. These observations suggest that, since the SAT has very little effect on who is in a college's freshman class, the SAT will also have very little effect on any college outcome.

EFFECT ON ADMISSIONS

College officials mean several different things when they say that they want to maximize their correct admissions decisions. On the one hand, they want to admit students who will make satisfactory freshman grades and go on to earn degrees and to reject those who will not finish their programs. But college admissions officers also care about the kinds of mistakes they make; they want to minimize the number of potentially unsuccessful applicants they admit and minimize the rejection of applicants who would have succeeded had they been admitted.

It is also important to note that officials from different colleges value different outcomes of admissions decisions, just as they place different values on various admissions criteria. Therefore, we should employ a variety of standards in evaluating the correctness of admissions decisions. However, since an admissions policy that uses high school rank alone makes the same admissions decisions for 90% of the applicants as a policy that also includes the SAT (i.e., the SAT is statistically redundant), the differences between the two policies should be minimal.

Correct admissions forecasts increase acceptances of applicants who earn satisfactory freshman grades and eventually graduate. Correct forecasts also increase the rejection of applicants who would earn unsatisfactory freshman grades and eventually drop out. Given a college that requires a 2.5 predicted freshman GPA for admission and a similar criterion of success for the actual freshman GPA, my study of the NLS data shows that such a college will make only 2.7 more correct admissions forecasts per hundred applicants when it uses the SAT in addition to high school rank than when it uses high school rank alone.[9]

The results are similar when a 3.0 predicted freshman GPA is required for admission. With this higher admissions standard, basing predictions on both high school rank and SAT improves correct admissions forecasts by only 2.2 per hundred.

When the yardstick used for success is college completion, the predictive value of the SAT is even less impressive. Colleges now make only trivial improvements in their ability to forecast college completion correctly when they use the SAT and high school rank together instead of high school rank alone. Indeed, correct forecasts increase by only 0.1 per hundred when using the SAT and high school rank in conjunction with a 2.5 predicted GPA admissions standard and by only 0.2 per hundred when the admissions standard is raised to a predicted GPA of 3.0.

Both rank-alone and rank-plus-SAT admissions policies decrease the number of false positives (acceptances of students who will fail) and of false negatives (rejections of students who will succeed) below the level of chance. But the added predictive value of the SAT over high school rank alone in the NLS data is never more than 1.3 fewer errors per hundred admissions decisions.

In all, the additional correct decisions and the reductions in errors from using both the SAT and high school rank rarely exceed two to three decisions per hundred. This is because the SAT has little effect on who is in a college's freshman class. The major difference, therefore, between the two policies depends on the relatively few applicants for whom the two policies lead to differ-

ent admissions decisions. This difference is always small.

With a 2.5 predicted freshman GPA as an admissions standard, the rank-plus-SAT policy admits an additional 1.3% of applicants who would be successful, and the high-school-rank policy admits an additional 1.4% who would be successful. The comparable values for a predicted 3.0 admissions standard are 6.3% for the rank-plus-SAT policy and 4.6% for the high-school-rank policy.

Colleges are also concerned with improving the academic quality of their student bodies and increasing the number of years of higher education that their students complete. However, they cannot expect to achieve much improvement in these areas from adding the SAT to an admissions policy that includes high school rank.

For example, if a college had an applicant pool with SAT scores and high school ranks comparable to those in the NLS survey and if the college admitted applicants randomly or admitted them all, the result would be freshman grades averaging 2.83 and years of education averaging 15.82. Using a high-school-rank admissions policy with a predicted freshman GPA of 2.5 as an admissions standard increases the freshman grades to an average of 2.91 and the years of education completed to 15.92; a rank-plus-SAT policy increases these figures to 2.93 and 15.93 respectively. The

TABLE 1.

SAT Scores and High School Class Rank as Predictors of College Outcomes

Probability of Average Grades of C or Better

Class Rank %	SAT Total Score 500	700	900	1,100	1,300	1,500
10	.31	.44	.57	.70	.80	.88
30	.43	.57	.69	.80	.88	.94
50	.56	.69	.80	.88	.93	.97
70	.69	.80	.88	.93	.97	.99
90	.79	.88	.93	.97	.98	.99

Percentile Rank in Freshman Class

Class Rank %	SAT Total Score 500	700	900	1,100	1,300	1,500
10	5	8	14	21	31	42
30	8	13	21	31	42	53
50	13	21	30	41	53	65
70	21	30	41	53	64	75
90	30	41	52	64	75	83

Probability of Graduation

Class Rank %	SAT Total Score 500	700	900	1,100	1,300	1,500
10	.42	.48	.53	.58	.64	.69
30	.49	.54	.59	.65	.70	.75
50	.55	.60	.66	.71	.76	.81
70	.61	.67	.72	.77	.82	.88
90	.68	.73	.78	.84	.89	.94

TABLE 2.

Cross-Tabulation of Predicted Freshman Grade-Point Averages of 2,781 NLS High School Seniors at the College They Applied to as Their First Choice

Freshman Grade-Point Average Predicted From High School Rank	Freshman Grade-Point Average Predicted from High School Rank and SAT																					
	1.6	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5	2.6	2.7	2.8	2.9	3.0	3.1	3.2	3.3	3.4	3.5	3.6	N
1.9		3	1	6	2	3	1															16
2.0	1	1	4	6	4	6	2											N = 81				24
2.1			2	5	9	8	7	11	3			1						2.9%				46
2.2				3	6	10	25	14	11	7	2	1										79
2.3		N = 463			4	11	18	19	17	9	4	1										83
2.4		16.6%			2	5	21	22	40	27	7	2	2									128
2.5				1		1	7	23	31	44	43	13	5									168
2.6					1	2	6	12	26	37	46	40	13	5								188
2.7								6	19	21	54	58	42	21	8	1			N = 2,062			230
2.8								1	8	19	36	57	72	65	23	3	3		74.2%			287
2.9				N = 175					2	11	29	64	83	98	60	31	8	4				390
3.0				6.3%					1	3	9	19	64	124	126	92	51	11	1			501
3.1												12	23	43	104	153	144	93	31	12		615
3.2													1	1	1	3	6	5	5	2	2	26
N	1	4	7	21	28	46	87	108	158	178	230	268	305	357	322	283	212	113	37	14	2	2,781

A policy using high school rank and SAT scores will make predictions that are very similar to a policy using high school rank alone.

gain from adding the SAT is, therefore, an increase in average freshman grades of only .02 on a four-point scale and an increase in average years of educational attainment of only .01 years.

The advantages of using the SAT are slightly greater with a 3.0 admissions standard. Average freshman grades increase by .03 and average years of education by .05.[10]

THE VALIDITY STUDY SERVICE

In 1964 the College Board established the Validity Study Service (VSS) to help member colleges perform their own validity studies. In addition to such sources of national data as the NLS, summaries of VSS studies done for individual colleges are a valuable source of information about the validity of the SAT.

The College Board does not release VSS data on individual colleges or individual students. But it does make available summary information from all participating colleges, as well as information on subsamples of colleges.

If my argument about the predictive usefulness of the SAT is correct, the VSS data should also demonstrate that an admissions policy based on high school rank plus the SAT should predict freshman GPAs that correlate highly with those predicted by rank alone. In fact, these correlations are quite high: .868 for the average college that took part in the VSS between 1976 and 1980, .801 for the most selective colleges, and .840 for the least selective colleges. The correlations drawn from the VSS data are not significantly different from the correlation of .881 for applicants in the NLS data.

These figures strongly support my general conclusion that, when admissions are based on predicted freshman grades, a policy using high school rank and SAT scores will make predictions that are very similar to a policy using high school rank alone. Thus, as I argued above, the SAT will have little impact on who is in a college's freshman class, and it is, therefore, statistically redundant.

Using the VSS data, the average college could expect to increase freshman grades through a rank-plus-SAT admissions policy by a maximum of only .15 points on a four-point scale. In order to achieve this .15 gain, the college must admit only 20% of its applicants. Using a more realistic selection ratio of 80%, the mean gain in average freshman grades would be between .02 and .03, virtually identical to the range I estimated using the NLS data. The results are similar for the most and least selective colleges.

Recently, ETS has reported an SAT coefficient that is 25% larger than the one I derived from the NLS survey.[11] It comes from 191 colleges that used SAT scores and high school GPAs in validity studies for entering classes from 1977 to 1981. Despite this larger effect of the SAT, these new data from ETS confirm every aspect of my analysis.

For example, the new ETS data demonstrate that an admissions policy based on high school grades plus SAT scores generates predicted freshman GPAs that correlate highly with those based on high school grades alone. The correlation in the new ETS data is .881, virtually identical to the values I presented above for the NLS and for other VSS data. So the SAT is also statistically redundant in the new ETS validity data.

Moreover, the estimated increment in average freshman grades of the accepted applicants at colleges using the SAT is also small. The new ETS data show that the typical college from these 191 colleges improves the average freshman grades of its accepted students by only .022 on a four-point scale if it uses a rank-plus-SAT admissions policy rather than rank alone. Once again, this value is within the .02 to .03 range that I estimated from other VSS data and from the NLS.

ETS is fond of pointing out that the size of the relationship between SAT scores and freshman grades is often dramatically underestimated when the range of SAT scores is restricted.[12] In the extreme case, for example, if a college enrolls students whose SAT scores are all the same, SAT scores can show no relationship to freshman grades. To get around this problem, ETS often corrects for restricted range in its estimates of the validity of the SAT.

Recently, ETS searched its VSS records for the College Board and found 21 colleges in which the distributions of SAT scores and high school rank are virtually identical to those of the total SAT-taking population. In these carefully chosen colleges with unrestricted ranges for high school records and SAT scores, the optimal equation for predicting freshman grades using high school record and SAT together is better than any other equation I have ever seen. The multiple correlation is .65, and the unstandardized coefficient for the SAT is 10% larger than the one for the new ETS data I reported above. If any data should show large benefits of the SAT, it should definitely be these.

Yet they do not. Robert Cameron, executive director for research and development at the College Board, gave me these data to analyze. I find no less statistical redundancy in these data than in any others for predicting freshman grades using the rank-plus-SAT admissions policy or the rank-alone policy. The correlation is .889, which is nearly identical to the .881 for the new ETS data. Furthermore, the gains in freshman grades of students selected by the rank-plus-SAT policy average only .03 higher on a four-point scale – again, virtually the same gains I reported above for the NLS and the VSS.

For all practical purposes, then, all these data confirm my conclusions. An admissions policy based on a combination of applicants' high school records and SAT scores admits and rejects nearly the same people as one based on high school records alone. Thus colleges that admit applicants according to their academic eligibility now benefit little from adding the SAT to their admissions requirements.

DEFENDING THE SAT

From a practical viewpoint, most colleges could ignore their applicants' SAT scores without appreciably altering the overall accuracy of their admissions decisions. This is contrary to the advice offered to colleges by the College Board and ETS that they should use their applicants' SAT score reports to make better academic selection decisions from their pools of applicants. The College Board and ETS might argue that, even though most colleges could ignore their applicants' SAT scores without appreciably altering their admissions outcomes, they would nonetheless sacrifice small but worthwhile gains in student quality.

If there were no costs involved in using the SAT, it could reasonably be argued that a massive admissions testing program is a good thing because it harms no one yet helps colleges make slightly better decisions. But the SAT has substantial costs that, in my opinion, outweigh by far the minimal help that colleges derive from using it.

Many of these costs are obvious. High school students who plan to attend college spend at least $30 taking the SAT one or more times and the Preliminary Scholastic Aptitude Test once. Many students also pay as much as $600 for commercial coaching aimed at improving their test scores. More important, many high schools use valuable instructional time on courses that prepare students for the SAT and other similar tests. In these courses, students hurriedly memorize long lists of vocabulary words and practice answering SAT-type questions from sample tests. This time and effort might be better spent learning academic subjects.

But the largest costs of the SAT may not be the monetary ones. These costs involve the anxiety of high school students and their parents, who nervously anticipate both the test itself and the arrival of the scores. A 1977 survey prepared for the College Board and designed to monitor the attitudes and reactions of students who take the SAT shows that nearly one-fifth of the students are "very nervous" about taking the SAT, while another 56% are "a little nervous."[13] Applying these figures to the entire population of students who take the SAT means that more than three-quarters of a million students faced the SAT with some trepidation in 1982-83.

Defenders of the SAT often argue that college-bound students can use their SAT scores to help them select a college appropriate for their level of ability.

"Whatever became of auto repair and bookkeeping?"

ETS encourages students to use their score reports in this way. No doubt most students do want to apply to colleges where they have a reasonable chance of being admitted and of doing satisfactory work. Students seldom apply to Ivy League schools, for example, unless they have good grades and high SAT scores. This kind of self-selection also seems to take place at other colleges.

However, neither the College Board nor ETS has ever tried to find out how prospective applicants actually use their SAT score reports. They have not tried, for example, to estimate the improvement in preselection of academically appropriate colleges that could be attributed to students' use of their SAT scores – that is, how much more accurate their preselections would be than if they used only their high school grades. In fact, the 1977 survey of student attitudes showed that "the majority of students think that one's performance on the SAT has little to do with how well one will do in college."[14]

Defenders of the SAT claim that the test has helped to guarantee equal opportunity. According to ETS,

> selective admission to higher education was far more a matter of class and economic status prior to the use of national admissions tests than it has been since. The introduction of tests resulted in a substantial increase in opportunities for educational advancement of low-income students by providing a credible demonstration that many such students from schools without reputations for educational excellence could succeed in the demanding academic programs of the most selective institutions.[15]

This version of the history of education is vastly oversimplified. To begin with, many elite colleges have *always* used admissions tests that allowed applicants from "schools without reputations for educational excellence" to demonstrate their competence. It is true that these examinations measured mastery of traditional academic subjects, including ancient languages, and that they therefore put students who had not studied these subjects at an overwhelming disadvantage. And this remained the case when, for reasons of administrative convenience, the leading selective colleges in the U.S. banded together early in this century to establish a single set of admissions tests.

It is also true that the SAT was introduced in 1926 to help selective colleges identify students who did poorly

Colleges that admit applicants according to their academic eligibility now benefit little from adding the SAT, as well.

on existing achievement tests because of their poor academic preparation but who also showed promise of doing good academic work if admitted to a good college. However, ETS has never tried to determine whether the SAT has achieved this objective. ETS has not, in other words, tried to determine whether students from secondary schools "without a reputation for educational excellence" are at a greater advantage when taking the SAT than they are when taking achievement tests.

But even if the SAT *has* reduced the handicap associated with coming from an obscure school, it hardly follows that the SAT has reduced the handicap associated with coming from a poor family. Just as predominantly upper-middle-class schools usually enroll a minority of students from the working class, so too predominantly working-class high schools usually enroll a minority of upper-middle-class students. Altering college admissions policies to admit more students from less well-known high schools may simply create opportunities for such upper-middle-class students living in less-privileged communities. Such a policy could increase rather than reduce the strength of the association between socioeconomic background and college admission.

It is tempting to try to resolve these uncertainties with historical evidence. Robert Mare, for example, has examined the impact of socioeconomic background on whether U.S. men who finished high school between about 1925 and 1969 went to college.[16] Using data from the 1973 replication of the Occupational Changes in a Generation survey, Mare showed that the impact of socioeconomic background on college entrance actually *increased* during these years. This is not to say that colleges grew more exclusive. On the contrary, college students became far more socially and economically heterogeneous.

But the democratization of higher education was neither as rapid nor as complete as the democratization of secondary education. Among those who finished high school, therefore, socioeconomic background became a somewhat better predictor of college attendance over the period Mare studied.

It would, however, be a mistake to rely on evidence of this kind to assess the impact of tests such as the SAT. Since virtually all high school graduates have always been able to get into *some* college regardless of their SAT scores, high school record, or other factors, the introduction of the SAT is most likely to affect *where* students go to college rather than *whether* they go. I know of no evidence regarding changes in the social and economic composition of the student population at highly selective colleges since they began relying on the SAT. Even if I had such data, I suspect that other factors, such as changes in the ratio of college costs to median family income, changes in the availability of college scholarships and loans, and, perhaps most important, changes in the pattern of recruiting activity at elite colleges, have affected the composition of the student bodies at these colleges more than the SAT has.

Proponents of the SAT often defend it by pointing out the problems with the obvious alternatives to it. Eliminating all external standardized testing, for example, in favor of dependence on applicants' high school records alone might tempt high schools to lower their grading standards. According to this argument, if a high school is under pressure from parents to place its graduates in selective colleges and if the school knows that such admissions depend on the grades it gives students, the school may be tempted to give higher grades. And if one school lowers standards, then others will feel compelled to follow suit in order to maintain their competitive position.

I maintain, however, that colleges can minimize this problem either by paying more attention to class rank or by demanding that high schools supply grade distributions for their entire graduating classes. Nevertheless, an external examining system may allow colleges to compensate more fully for differences in high school grading standards. But it does not follow that such an examining system ought to include the SAT.

Defenders of the SAT point out that replacing it with conventional achievement tests in the subject areas of the secondary school curriculum could lead the U.S. to a system not unlike the British O-level and A-level exams and thus encourage a higher degree of curricular uniformity in the high schools than now exists. According to this argument, such a system might not only reduce the variety in students' choice of subjects but also limit access to higher education for poor and minority students who might not have chosen early in their schooling to follow a precollegiate curriculum.

Certainly each of these arguments needs to be carefully evaluated before we can decide whether the SAT is, on balance, a good thing. My argument is only that the use colleges make of their applicants' SAT score reports does not now help them improve the accuracy of their admissions decisions in terms of the academic performance or graduation rates of the students they admit. Thus, if the College Board is to keep urging its member colleges to require applicants to submit SAT scores, it should do so for reasons other than the predictive value of the test. If other reasons for giving the SAT exist, such as helping students make better decisions or giving students a chance to be measured outside their high schools, the College Board owes it to all concerned to do more than merely offer these reasons. It should also demonstrate why the reasons are valid.

---

1. The author wishes to thank Karl Alexander, Albert Beaton, Carol Blumberg, Norman Bowie, George Brabner, Marvin Brams, Ken Eckhardt, Art Hoerl, Saul Hoffman, Christopher Jencks, Arthur Jensen, Robert Linn, Victor Martuza, Samuel Messick, Lou Mosberg, Peter Mueser, Michael Olneck, David Owen, Leonard Ramist, Dale Trusheim, and other participants in the ETS seminar for their criticisms of early drafts.

2. James Crouse, "Does the SAT Help Colleges Make Better Selection Decisions?," *Harvard Educational Review*, May 1985, pp. 195-219. The August 1985 issue of the same journal carried a response from the president of the College Board, George Hanford, "Yes, the SAT Does Help Colleges," pp. 324-31. My response to Hanford appeared in the November 1985 issue.

3. College Board, *Taking the SAT: A Guide to the Scholastic Aptitude Test and the Test of Standard Written English* (New York: College Entrance Examination Board, 1978), p. 48.

4. Educational Testing Service, *Test Use and Validity: A Response to Charges in the Nader/Nairn Report on ETS* (Princeton, N.J.: ETS, 1980), p. 5.

5. *Undergraduate Admissions: The Realities of Institutional Policies, Practices, and Procedures* (New York: College Entrance Examination Board, 1980), p. 14.

6. See, for example, Susan F. Ford and Sandy Campos, *Summary of Validity Data from the Admissions Testing Program Validity Study Service* (Princeton, N.J.: College Board, 1977); and *Test Use and Validity. . . .*, p. 5.

7. For information about the Project Talent Survey, conducted initially in 1960, see John C. Flanagan et al., *The American High-School Student* (Pittsburgh: University of Pittsburgh Press, 1964); for information about the National Longitudinal Study, conducted initially in 1970, see R. Paul Moore, *National Longitudinal Study Base Year, First, Second, and Third Follow-Up Data File Users Manual* (Research Triangle, N.C.: Research Triangle Institute, 1978); and for information about the 1980-82 High School and Beyond Survey, see *High School and Beyond 1980 Cohort First Follow-Up (1982) Data File User's Manual* (Washington, D.C.: National Center for Education Statistics, 1983).

8. See, for example, Warner V. Slack and Douglas Porter, "The Scholastic Aptitude Test: A Critical Appraisal," *Harvard Educational Review*, May 1980, pp. 154-75; and Allan Nairn et al., *The Reign of ETS: The Corporation That Makes Up Minds* (Washington, D.C.: Ralph Nader, 1980).

9. This amounts to a *proportional* increase of 29.3% more correct admissions decisions. Nonetheless, I doubt that most colleges would notice such a small actual increase, no matter what the proportional improvement. In my judgment, the problem with proportional improvements is that they can seem large just because the base from which the improvement is calculated is small. ETS often uses proportional improvements in response to its critics who point out that actual improvements in predictions are small. See, for example, Rex Jackson, "The Scholastic Aptitude Test: A Response to Slack and Porter's 'Critical Appraisal,'" *Harvard Educational Review*, November 1980, pp. 382-91. Indeed, ETS has accused others of denigrating the usefulness of the SAT by failing to stress proportional improvements. See Warren Willingham and Leonard Ramist, "The SAT Debate: Do Trusheim and Crouse Add Useful Information?," *Phi Delta Kappan*, November 1982, pp. 207-8. Yet I have never seen a discussion from ETS about whether proportional improvements give a misleading picture of the importance of absolute gains from using the SAT.

10. The differences in the two admissions policies depend on the small number of students who are treated differently when one admissions policy is used instead of the other. Thus, with the lower admissions standard, 2.9% of all applicants are admitted with the rank-plus-SAT policy who would not be admitted with the rank-only policy; 5.4% of applicants are admitted with the high-school-rank policy who would not have been admitted with the rank-plus-SAT policy. The average freshman grades of these two groups differ by only .14 points (2.60 - 2.46) on a four-point scale. Their average years of education differ by only .07 years (15.55 - 15.48). The differences are likewise small when the higher standard is used.

11. Thomas F. Donlon, ed., *The College Board Technical Handbook for the Scholastic Aptitude Test and Achievement Tests* (New York: College Entrance Examination Board, 1984), fn. on p. 146. Actually, the regression coefficient value presented there is incorrect. The correct value is the cited value divided by 100.

12. Ibid., pp. 144-46.

13. *SAT Monitor Program* (Princeton, N.J.: Response Analysis Corporation, July 1978), p. 22.

14. Ibid., p. 37.

15. *Test Scores and Family Income: A Response to Charges in the Nader/Nairn Report on ETS* (Princeton, N.J.: ETS, 1980), p. 10.

16. Robert Mare, "Change and Stability in Educational Stratification," *American Sociological Review*, vol. 46, 1981, pp. 72-87. K

EA8680-1A

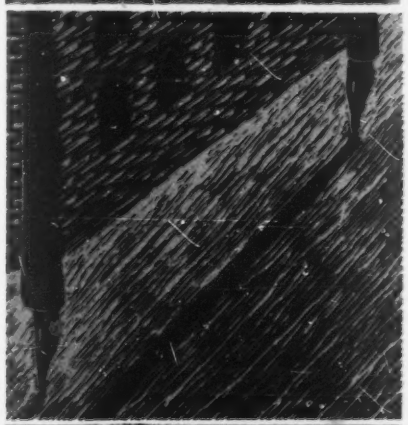
PHI DELTA KAPPAN APRIL 1986

KAPPAN

The Changing Face Of Teachers

Kottkamp · Provenzo · Cohn
Dan C. Lortie

State-Sponsored School Improvement
— A Special Section



Phi Delta April 1986 • Volume 67 • Number 8

KAPPAN

FEATURES

559 **Stability and Change in a Profession: Two Decades of Teacher Attitudes, 1964-1984,** *by Robert B. Kottkamp, Eugene F. Provenzo, Jr., and Marilyn M. Cohn*

568 **Teacher Status in Dade County: A Case of Structural Strain?,** *by Dan C. Lortie*

576 **Well Below His Ability,** *by James A. Mecklenburger*

State-Sponsored School Improvement: A Special Section

578 **State Initiatives Can Foster School Improvement,** *by Beverly Anderson and Allan Odden*

582 **How Successful State Education Improvement Programs Work,** *by Allan Odden and Beverly Anderson*

586 **State-Sponsored Schoolwide Improvement Programs: What's Going On in the Schools?,** *by Eleanor Farrar and Patricia Flakus-Mosqueda*

590 **How State Instructional Improvement Programs Affect Teachers and Principals,** *by Alan Davis and Allan Odden*

594 **Improving Schools: The State Role,** *by Susan Fuhrman, Eugene Huddle and Jane Armstrong*

597 **Chapter 2: For Better or Worse?,** *by Anne T. Henderson*

602 **Writing for Publication: Playing to Win,** *by Kenneth T. Henson*

605 **An Alternative Form of Collective Bargaining,** *by Patricia Smith and Russell Baker*

607 **A Letter from England: Professionalism and Education,** *by Richard de Zoysa*

DEPARTMENTS

554 **THE EDITOR'S PAGE,** Getting to Bedrock

555 **WASHINGTON REPORT,** *by Anne C. Lewis*

557 **STATELINE,** *by Chris Pipho*

609 **PAR,** Student Promotion and Retention, *by Monica Overman*

613 **DE JURE,** Establishment Clause Case Reveals Flaws in the Court System, *by Thomas J. Flygare*

614 **PROTOTYPES,** Do Schools Discipline Students Too Much?, *by Robert E. Rubinstein*

615 **BOOKS,** *Laurence D. Brown*

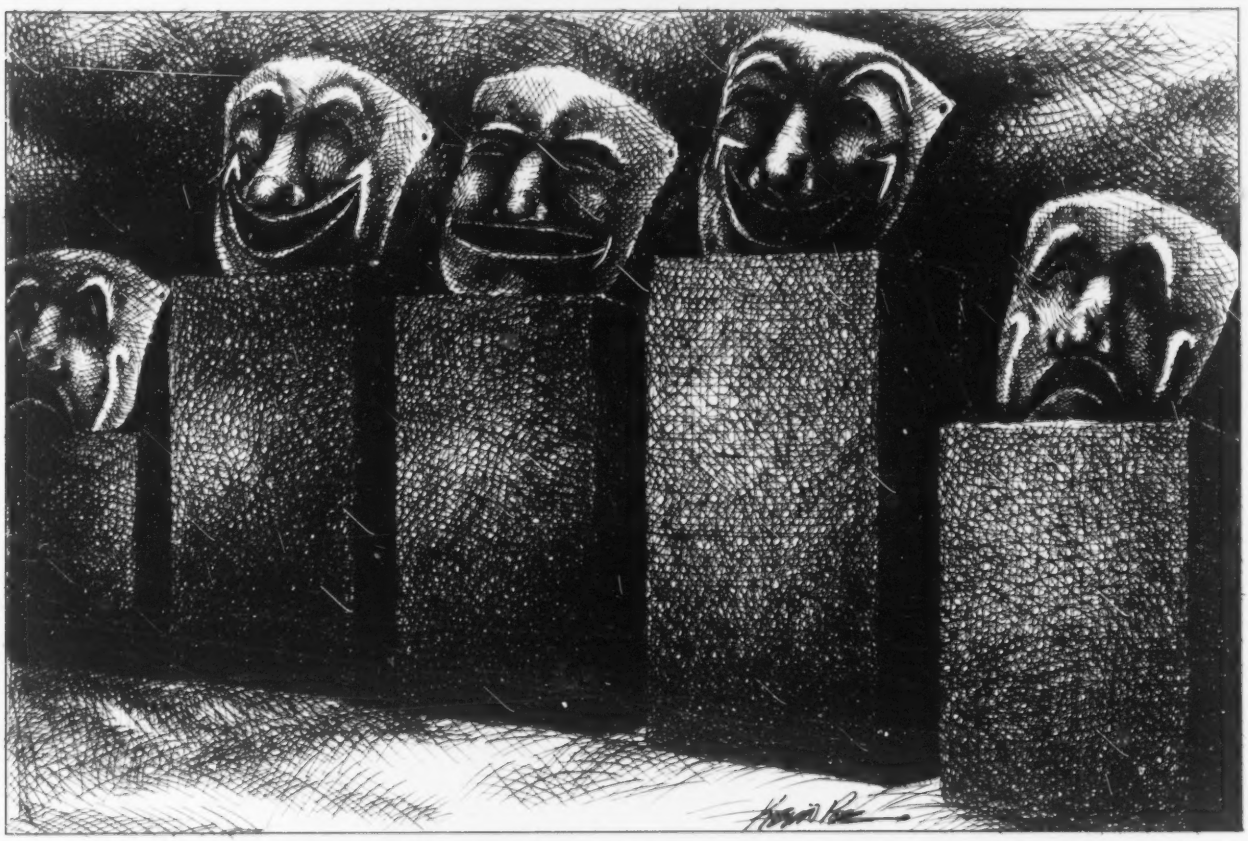
620 **NEWSNOTES**

623 **BACKTALK**

Cover Illustration: Dave Cutler

Stability and Change In a Profession: Two Decades of Teacher Attitudes, 1964-1984

by Robert B. Kottkamp, Eugene F. Provenzo, Jr., and Marilyn M. Cohn



ROBERT B. KOTTKAMP (Rutgers University Chapter) is an assistant professor of educational administration at Rutgers University, New Brunswick, N.J. EUGENE F. PROVENZO, JR. (University of Miami Chapter) is a professor of foundations of education at the University of Miami, Coral Gables, Fla. MARILYN M. COHN (Washington University Chapter) is director of teacher education at Washington University, St. Louis. *The research on which this article is based was supported by the National Institute of Education (Contract No. NIE-G-83-0067) and by grants from the University of Miami and Rutgers University Chapters of Phi Delta Kappa. The opinions expressed in this article are those of the authors and do not reflect the position, policy, or endorsement of the funding agencies. The authors wish to acknowledge the assistance of Dan C. Lortie, who made available his 1964 data; Leonard Britton, superintendent of the Dade County (Fla.) Public Schools; Pat Tornillo, executive vice president of the United Teachers of Dade (American Federation of Teachers); and the teachers and administrators of the Dade County schools. The comments of Jeffrey K. Smith of Rutgers University on an earlier draft were also helpful.*

IN THIS ERA of intense public concern with education and of pressure for education reform, the viewpoints of classroom teachers have

Dan Lortie's 1964 study, *Schoolteacher*, provided a valuable landmark for subsequent examination of the teaching profession. Now three researchers have replicated Mr. Lortie's work and provided us with an important update on the views of teachers.

Illustration by Kevin Pope

too often been ignored. When teacher opinions *have* been surveyed, the reports of those surveys have failed to provide a historical perspective on the views of teachers.

Because it presents data on stability and change in teacher attitudes over the 20-year period from 1964 to 1984, this article provides the previously missing historical perspective. Our observations spring from questionnaires completed by more than 6,500 Dade County (Florida) teachers in 1964 and more than 2,700 Dade County teachers in 1984, from intensive interviews with approximately 100 of the teachers in 1984, and from our analysis of the major forces that have affected teachers and their work over the past 20 years.

One aim of our study was to understand how major social forces are related to shifts in teacher attitudes and orientations toward work. This article, which is the first report of our study, contains only the frequency distributions of responses to survey questions asked of Dade County teachers in 1964 and again in 1984. As such, it provides a purely descriptive analysis of historical shifts in teacher attitudes. The historical and interview data, which are only just beginning to be analyzed, should provide a rich explanatory base for some of the changes we describe here. We believe it is important to present the following data at this time because of the questions they raise about the current debate on education reform.

The baseline survey data were collected in 1964 from Dade County public school teachers by Dan Lortie for his seminal book, *Schoolteacher*.[1] Despite enrollment pressures caused by the "baby boom," 1964 was a year of relative calm in U.S. education. The Russians had put Sputnik I into orbit seven years earlier; the Vietnam war, the women's movement, the back-to-basics movement, and accountability would not have major effects on the schools until some years later. In Dade County, Lortie's survey also predated the major movement for unionization, the first teacher strike, the implementation of desegregation plans, and several of the waves of immigrants from Cuba, South America, and the Caribbean.

Lortie chose Dade County as the site for his research in part because the countywide school system provided a rich diversity of community and school types – and in part because the Miami area was populated by persons born, and presumably reared, in all sections of the U.S. Lacking the resources to gather a national sample, Lortie reasoned that Dade County was a better-than-average setting for ascertaining the attitudes of teachers with varying regional and cultural origins. His data were collected on a single day and represented virtually the entire Dade County teaching population, minus those who were absent from school (N=6,539).[2]

The 1984 survey data were collected from teachers in all 251 Dade County schools. We drew a 40% stratified random sample of classroom teachers from each school. (The sample was stratified by school to allow us to examine issues at the school level.) Principals and union stewards were asked to encourage teachers to return the questionnaires, which were distributed through the school mail system. Follow-up efforts included mailings and phone calls to schools and homes. These procedures yielded a 64% overall response rate (N=2,718).

COMPARING THE STUDIES

During the 20 years that elapsed between Lortie's study and ours, the teaching profession changed demographically. In addition to looking at these demographic changes, we will examine stability and change in teacher attitudes and views on several important educational issues. Finally, we will examine the relationship between the empirical findings about transitions in the profession and the recent critiques of teaching and proposed reforms put forth by the various national commissions and task forces.

TABLE 1.
Age of Dade County Teachers

	1964 %	1984 %
20-29	30.2	10.7
30-39	30.7	35.2
40-49	22.0	27.4
50-59	17.0	20.2
60 or over	.2	6.5

DEMOGRAPHIC CHANGES

Demographically, the teaching force of 1984 differs from that of 1964. Teachers are older and more experienced today than in the past. In 1964 only 39% of Dade County teachers were over 40 years of age, and only 43% had more than 10 years of total teaching experience. Today, 54% of the teaching force is 40 or older, and 67% of teachers have taught for more than 10 years (see Tables 1 and 2).

Teachers also have more formal education today. In 1964 only 25% of the teaching force possessed at least a master's degree; 20 years later 58% had completed that much schooling (Table 3).

Finally, the ethnic composition of the teaching force has changed considerably (Table 4). In his 1964 study Lortie used only three categories: White, Negro, and Other. In Dade County, the percentage of teachers who are white has declined from 83% to 59% since 1964, the percentage who are black has grown from 17% to 25%, and a totally new Hispanic segment (16% of the total group) appeared in the 1984 survey.

Although their effects cannot be isolated, numerous social forces have led to these changes in the demographic characteristics of teachers. Among the social forces are: a declining birthrate, which resulted in less demand for new

TABLE 2.
Teaching Experience of Dade County Teachers

	Years in Dade County Public Schools		Total Years Teaching	
	1964 %	1984 %	1964 %	1984 %
1 year or less	14.4	7.0	10.5	2.5
2-5 years	31.7	18.2	23.3	11.8
6-10 years	26.4	19.1	23.2	18.5
11-15 years	13.9	22.3	16.2	22.5
16-20 years	6.1	16.3	9.7	18.9
21-25 years	3.1	10.1	7.5	12.4
26 years or more	4.4	7.1	9.5	13.4

(i.e., young) teachers; stagnation and inflation in the economy, which made it more difficult for experienced teachers to find alternative employment; the tying of salary increments to formal education, which caused more teachers to earn graduate degrees; and the Cuban revolution, which brought large numbers of middle-class Hispanic immigrants into the Dade County teaching force.

Data on places of birth support Lortie's contention that Dade County is a "better-than-average site"[3] for ascertaining the attitudes of teachers with varying regional and cultural origins. Although some shifts in the population occurred between 1964 and 1984, the overall picture today is still one of great diversity (Table 5). The most noteworthy change is the large influx (13%) of Cuban-born teachers – mostly members of the Cuban middle class – who arrived in the U.S. in the wave of immigration following the Cuban revolution.*

Clearly, the Dade County teaching force changed significantly on a number of demographic variables between 1964 and 1984. Moreover, except for the large influx of Cuban-born teachers, the changes in Dade County are similar to those that have occurred nationwide.[4] Whether the trends toward greater age, experience, and formal education in part "explain" some of the shifts in teacher attitudes that we will chronicle in the next section of this article remains a question. Our larger study will seek to understand the relationships among demographic changes, other social forces, and attitudinal changes.

TEACHER ATTITUDES/PERCEPTIONS

Satisfaction with job and workplace. Currently there are major concerns about attracting and keeping good teachers in the workforce. Satisfaction with work and with the conditions of work are indices related to the likelihood that individuals will remain in their jobs. The Dade County teachers were asked in 1964 and in 1984 how satisfied they were with teaching as a job and with their schools as workplaces. Although overall satisfaction rates remained relatively high across that 20-year period, slight declines occurred in levels of teacher satisfaction with both job and workplace. The decline was more marked with regard to the school as a workplace than with regard to teaching as a job (Table 6).

In 1964, 81% of the Dade County teachers expressed some level of satisfaction with teaching as a job; in 1984, by contrast, only 76% did so. In 1964, 80% of the Dade County teachers expressed some level of satisfaction with their schools; in 1984 only 68% did so.

Teaching orientation. Public and professional discussions of educational goals in recent years have focused on moving "back to basics." How has this trend affected Dade County teachers? Asked to describe the emphasis in their own classrooms, teachers seemed to be

TABLE 3.
Formal Education of Dade County Teachers

	1964 %	1984 %
Less than bachelor's degree (1964 only)	.4	--
Bachelor's degree	11.1	8.1
Bachelor's plus courses	63.6	34.5
Master's degree	7.3	18.5
Master's plus courses	17.1	37.6
Doctorate	.4	1.4

TABLE 4.
Ethnic Composition of Dade County Teachers

	1964 %	1984 %
White non-Hispanic (White, 1964)	82.5	59.3
Black non-Hispanic (Negro, 1964)	17.2	24.4
Hispanic (1984 only)	--	15.6
Asian and Native American (1984 only)	--	.7
Other (1964 only)	.3	--

TABLE 5.
Birthplace of Dade County Teachers

	1964 %	1984 %
Dade County	12.5	16.9
Florida, outside Dade County	11.1	9.8
Md., Va., N.C., S.C., Ga., Ala., Miss., Tenn., Ky.	21.5	13.7
N.Y., N.J., Del., Penn.	24.7	24.2
New England	5.2	4.1
Wis., Mich., Minn., Ill., Ind., O., W.Va.	15.2	9.4
A state west of the Mississippi	6.9	4.9
Foreign country (1964 only)	2.9	--
Cuba (1984 only)	--	13.0
Other foreign country (1984 only)	--	4.0

TABLE 6.
Teachers' Feelings About Job and Workplace

	Feelings About My Job		Feelings About My School	
	1964 %	1984 %	1964 %	1984 %
Very satisfied	31.9	27.4	30.4	23.6
Satisfied	26.5	29.1	26.4	25.3
More satisfied than not	22.6	19.7	23.7	18.6
Equally satisfied and dissatisfied	10.0	10.7	11.6	12.4
More dissatisfied than satisfied	6.0	7.6	5.4	9.9
Dissatisfied	1.9	3.5	1.6	5.6
Very dissatisfied	1.0	2.0	.9	4.6

*To ascertain whether the 1984 data on teacher attitudes had been significantly affected by the sizable population of Cuban-born teachers (which did not exist in Dade County in 1964), we ran analyses of the 1984 data using both the total teacher sample and a subsample composed of only native-born teachers. Similarly, because the age of the teaching force had changed markedly over the 20-year interval, we ran separate analyses of the 1964 data and of the 1984 data from two age groups – 30-39 and 50-59 – to see whether response patterns similar to those of the total teacher samples would obtain. In all three of these additional analyses, we found no major differences in the patterns of stability and change from those that we have reported in this article. All patterns reported in this article are based on the complete data set for each year.

TABLE 7.
Dade County Teachers' Descriptions of Their Teaching Orientation

	1964 %	1984 %
I'm pretty much the "no-nonsense, get-the-learning-of-the-subject-matter-done" kind of teacher.	5.7	15.5
I tend toward the subject-matter emphasis, but I think other things are important too.	73.7	67.0
I'm about 50/50 on this.	17.6	16.3
I tend away from emphasis on subject matter, as I consider other things more important.	2.5	1.0
I think that emphasis on subject matter is the mark of a poor teacher.	.4	.1

TABLE 8.
Dade County Teachers' Perceptions of the Most Important School Goals for Younger and Older Students

	Most Important Goal for Students (K-6)		Most Important Goal for Students (7-12)	
	1964 %	1984 %	1964 %	1984 %
The basic tools for acquiring and communicating knowledge — the three R's	27.1	42.0	1.5	2.4
Efficient *use* of the three R's — the basic tools for acquiring and accumulating knowledge	44.5	35.1	14.0	24.3
A continuing desire for knowledge, the inquiring mind	9.3	8.5	29.5	31.0
The ability to live and work with others	8.8	5.4	10.4	5.8
The habit of weighing facts and imaginatively applying them to the solution of problems	2.7	1.2	17.0	13.4
A sense of right and wrong — a moral standard of behavior	2.0	4.9	3.0	5.0
An understanding of the government and a sense of civic responsibility	.1	0	1.4	1.2
An emotionally stable person, prepared for life's realities	4.4	2.1	16.1	10.5
Loyalty to America and the American way of life	.6	.2	1.8	.8
Information and guidance for a wise occupational choice	.3	.1	4.8	4.6
A well-cared-for, well-developed body	--	.1	.2	.2
Other	--	.4	.3	.8

TABLE 9.
The Kinds of Students Whom Dade County Teachers Prefer to Teach

	1964 %	1984 %
Students whose emotional needs are a challenge to the teacher	6.0	3.0
"Nice kids," from average homes, who are respectful and hard-working	36.2	46.6
Creative and intellectually demanding students who call for special effort	40.0	27.9
Underprivileged students from difficult or deprived homes for whom school can be a major opportunity	12.2	13.5
Students of limited ability (sometimes called "slow learners") who need unusual patience or sympathy	5.7	9.0

more subject-matter-oriented in 1984 than their counterparts had been in 1964. The percentage who described their orientation toward teaching as "no-nonsense, get-the-learning-of-the-subject-matter-done" nearly tripled, moving from 6% to 16% (Table 7).

Goals of schooling. Teachers' perceptions of the goals of schooling are related to their orientations toward teaching. In 1964 Lortie asked Dade County teachers to choose the "most important thing" schools should do for elementary (K-6) and for secondary (7-12) students — a question that we repeated in our 1984 study. It is important to note that all the teachers in each sample reported their perceptions of goals for both the elementary and the secondary levels, not just for the level at which they taught.

Generally speaking, the goals teachers chose in the two studies reveal a shift toward more rudimentary achievement for students today (Table 8). At the elementary level, a marked shift occurred from a focus on the efficient *use* of the three R's in 1964 to the simple *acquisition* of the three R's in 1984. At the secondary level, the largest shifts were away from intra- and interpersonal skills and creativity and toward efficient use of the three R's. If the acquisition of the three R's and the ability to use these skills are prerequisites for attaining many of the other goals listed as choices, the pattern of responses from 1964 to 1984 can be seen as a general lowering of teachers' expectations for their students.

The kinds of students whom teachers prefer to teach. In the years between 1964 and 1984 the student population changed in Dade County and in the U.S. as a whole. Students in 1964 were part of the "baby boom" generation. Today's student population includes larger proportions of blacks and Hispanics. Over the 20-year interval growth has also occurred in the number of students classified as eligible for various kinds of special education, and a corresponding increase has occurred in the number of teacher specialists trained to work with children who have learning disabilities and social problems.

In both 1964 and 1984, Dade County teachers were asked their first preference regarding the kinds of students to be placed in their classrooms for the coming year. The modal response of teachers to this question has changed over the 20-year period (Table 9). Twelve percent fewer teachers in 1984 chose the response that ranked highest

among teachers in 1964, chosen by 40% of them: "creative and intellectually demanding students who call for special effort." Ten percent more teachers in 1984 than in 1964 chose the most popular response in the 1984 study, chosen by 47% of the Dade County teachers: " 'nice kids,' from average homes, who are respectful and hard-working." (A number of teachers who chose the "nice kid" option underlined the word *respectful*, an indication of the importance they placed on this attribute.)

Three percent fewer teachers in 1984 than in 1964 said that they wished to teach students with challenging emotional needs, but 3% more teachers in 1984 than in 1964 said that they wished to teach youngsters with limited cognitive ability.

Overall, the Dade County teachers in 1984 were more likely than their 1964 counterparts to seek students who would not drain them physically, intellectually, and emotionally. We have not yet examined the relationship of this phenomenon to the aging of the teaching force.

Sources of help for teachers. The potential sources of help with curricular and pedagogical problems have changed somewhat over the years. A growing number of curricular and supervisory specialists are available to teachers who need assistance. More recently the professional literature has placed great emphasis on the importance of having principals serve as instructional leaders within their schools.

The sources from whom teachers receive help on curriculum and teaching methods have changed in important ways since 1964 (Table 10). Teachers were less likely in 1984 (38%) than in 1964 (54%) to say that they receive help from line administrators and supervisors (principals, assistant principals for curriculum, and department chairpersons). But teachers in 1984 were more likely than teachers in 1964 to receive help from resource teachers and subject-matter specialists and from "other" sources. Teacher colleagues in the school were the primary source of help – both in 1964 and in 1984. The second most important source of help in both years was the assistant principal for curriculum, though 7% fewer teachers listed this source in 1984 than in 1964.

Assessment of teaching effectiveness. A strong emphasis on accountability since the mid-1970s and several recent reports on U.S. education by national commissions and task forces have increased public pressures for more thorough assessment of teachers' effectiveness. In light of these pressures, have teachers altered their perceptions of their own effectiveness or the gauges by which they judge effectiveness? Two questions that Lortie asked Dade County teachers in 1964 provide some information on these issues: 1) How difficult is it to assess one's own effectiveness? 2) What indicator is a good teacher most likely to rely on in assessing his or her own effectiveness?

The modal response to the first question changed between 1964 and 1984 (Table 11). In 1964 the largest percentage of teachers (49%) said that it is possible to know one's own effectiveness *at times*. In 1984, by contrast, the largest percentage of teachers (52%) said that it is *relatively easy* to know when one is really teaching effectively. In other words, teachers see their own effectiveness as increasingly easy to ascertain.

Between 1964 and 1984, responses to the question of which indicator a good teacher is most likely to use in assessing his or her own effectiveness changed even more dramatically (Table 12). In 1964 the modal response (59%) was the teacher's perception of what students were learning compared to the teacher's conception of what they should be learning. In 1984 that response was still the most popular, but it was chosen by only 37% of the Dade County teachers – a difference of 22 percentage points.

TABLE 10.
Sources from Whom Dade County Teachers Receive the Most Help with Curriculum and Methods

	1964 %	1984 %
The principal	9.4	3.6
Central office supervisor(s)	4.5	1.6
Resource teachers/subject-matter specialists	.5	12.5
Assistant principal for curriculum/curriculum assistant	28.5	22.0
Department or grade-level chairperson	16.0	12.5
Other teachers in the school	36.8	36.0
Other	4.3	11.9

TABLE 11.
Dade County Teachers' Views of the Ease of Assessing Their Own Effectiveness

	1964 %	1984 %
It is *relatively easy* to know when one is really teaching effectively.	45.6	52.0
It is possible to know one's own effectiveness *at times*.	48.7	43.9
It is *relatively difficult* to know when one is really teaching effectively.	5.6	4.1

TABLE 12.
The Indicator That a Good Teacher Is Most Likely to Use in Judging Personal Effectiveness

	1964 %	1984 %
Reactions of other teachers familiar with their work and students	11.7	16.3
Opinions expressed by students generally	7.4	11.1
General observations of students in light of the teacher's conception of what should be learned	58.9	37.3
Assessments made by the principal	2.9	6.0
Assessments made by a "supervisor" or similar person	1.2	2.8
Results of objective examinations and various other tests	13.4	18.6
Reactions of students' parents	1.2	2.7
Other	3.3	5.3

More than a quarter of today's Dade County teachers claim to receive no satisfaction from extrinsic rewards.

The teachers in 1984 tended more often than their 1964 counterparts to rely on *external* indicators of effectiveness: other teachers, students, principals, supervisors, parents, and test scores. The indicators of effectiveness that increased in popularity the most between 1964 and 1984 were reliance on the reactions of other teachers and reliance on test scores. But the proportions of the teacher sample who cited the opinions of students, of the principal, of a supervisor, or of parents as the major gauge of teaching effectiveness have all more than doubled over the 20-year interval.

If we separate the indicators of effectiveness into two groups, *internal* and *external*, the change that occurred between 1964 and 1984 is a major one. In 1964, 59% of the Dade County sample believed that good teachers were most likely to gauge their effectiveness through self-assessment. In 1984, by contrast, 63% of the Dade County sample said that good teachers were most likely to rely on some type of external evaluation of effectiveness. (Remember, too, that more teachers in 1984 than in 1964 said that it is relatively easy to know when one is teaching effectively.)

Career and work rewards. In *Schoolteacher*, Lortie defined three categories of career and work rewards: extrinsic rewards, ancillary rewards, and psychic or intrinsic rewards. According to Lortie,

> The first includes what we usually think of as the "earnings" attached to a role and involves money income, a level of prestige, and power over others. . . . [S]ince they are experienced by all incumbents, they have an "objective" quality. Ancillary rewards are simultaneously objective and subjective; they refer to objective characteristics of the work which may be perceived as rewards by some (e.g., married women might consider the work schedules of teaching to be rewarding while men might not). Ancillary rewards tend to be stable through time and to be "taken for granted" rather than specified in contracts. . . . Psychic rewards consist entirely of subjective valuations made in the course of work engagement; their subjectivity means that they can vary from person to person.[5]

In his 1964 study Lortie asked Dade County teachers to specify the type of reward they saw as most important within each of the three categories. He also asked teachers to choose the most important category from *among* the three categories.

With regard to extrinsic rewards (Table 13), the largest proportion of Dade County teachers in 1964 – 37% – cited respect from others as the most important. Wielding influence followed closely, named by 36% of the sample. By 1984 the most frequent response of Dade County teachers was wielding influence, though 4% fewer teachers mentioned this reward in 1984 than in 1964. The second most popular response in 1984 was "I receive no satisfaction," mentioned by 28% of the sample – more than double its 1964 frequency (13%). In other words, more than a quarter of today's Dade County teachers claim to receive no satisfaction from extrinsic rewards.

Shifts have also occurred with respect to ancillary rewards (Table 14). In 1964 the most frequently chosen ancillary reward was "the special appropriateness of teaching for people like me," cited by 34% of the Dade County teachers. In 1984, however, the most popular response (cited by 35% of the teachers) was the schedule, which allows time for travel and family. The frequency of that response increased by 12 percentage points over the 20-year interval between Lortie's study and ours. Clearly, the ancillary reward that gives teachers the most satisfaction today is the ability to

TABLE 13.
Extrinsic Rewards That Are Most Satisfying To Dade County Teachers

	1964 %	1984 %
The salary I earn in my profession	14.3	14.2
The respect I receive from others	36.6	26.3
The opportunity to wield some influence	36.0	31.7
I receive no satisfaction at all from these things.	13.0	27.8

TABLE 14.
Ancillary Rewards That Are Most Satisfying To Dade County Teachers

	1964 %	1984 %
The relative security of income and position	22.6	16.1
The schedule (especially summer), which can permit travel, family activities, etc.	23.2	35.4
The opportunity that teaching offers to earn a living without much rivalry and competition with other people	4.8	2.8
The special appropriateness of teaching for people like me	34.1	28.6
None of these afford me satisfaction.	15.3	17.1

spend time *away* from work. The proportion of teachers who said they derive no satisfaction from ancillary rewards also increased slightly, from 15% in 1964 to 17% in 1984.

Dade County teachers have changed less with respect to the intrinsic or psychic rewards that give them the most satisfaction (Table 15). "Reaching" a student or a group of students was the modal response, both in 1964 and in 1984. The proportion of teachers who said they receive "no satisfaction" from intrinsic rewards was small – only 1.6% of the total sample in 1984, though this figure represents a fivefold increase over 1964. This change may be more important than the small numbers suggest, because teachers have always given the highest priority to the intrinsic rewards of their work.

Teachers in 1984 also differed from their counterparts in 1964 on which category of rewards – extrinsic, ancillary, or intrinsic – was most important to them (Table 16). The proportion of teachers who chose extrinsic rewards declined slightly, from 11.9% in 1964 to 11.3% in 1984. In both years, the largest number of teachers said that psychic or intrinsic rewards were the most important category. However, 6% less of the sample concurred on the importance of intrinsic rewards in 1984 than in 1964. Meanwhile, 6% more of the teachers in 1984 than in 1964 saw ancillary rewards as most important. This shift is an important indicator of change in the professional orientation of Dade County teachers.

The 20-year interval between 1964 and 1984 brought a gain of 8% in the number of teachers who favor differential pay and prestige.

TABLE 15.
Intrinsic Rewards That Are Most Satisfying To Dade County Teachers

	1964 %	1984 %
The opportunity that teaching gives me to study, read, and plan for classes	3.4	1.9
The chance that teaching offers to develop mastery of discipline and classroom management	1.1	1.4
The times I know I have "reached" a student or group of students and they have learned	86.2	86.7
The chance to associate with children or young people and to develop relationships with them	7.9	7.6
The chance that teaching gives me to associate with other teachers and educators	1.0	.7
I receive no satisfaction from these.	.3	1.6

TABLE 16.
Most Important Category of Rewards To Dade County Teachers

	1964 %	1984 %
The salary and respect received and the position of influence	11.9	11.3
The opportunities to study, plan, master classroom management, "reach" students, and associate with colleagues and children	76.3	70.2
The economic security, time, freedom from competition, and appropriateness for persons like me	11.8	18.4

TABLE 17.
Dade County Teachers' Attitudes Toward Differential Pay and Prestige

	1964 %	1984 %
Teachers who excel should receive extra income and prestige.	41.6	49.8
Teachers should receive more-or-less equal income and prestige.	58.4	50.2

Merit pay. In 1964 Lortie asked teachers whether or not they favored a system of differential pay and prestige, but his question did not use the words "merit pay." Fifty-eight percent of the teachers in his sample favored granting all teachers "more-or-less equal income and prestige" (Table 17). By 1984 the Dade County teachers who answered this same question split their responses almost evenly. Thus the 20-year interval between 1964 and 1984 brought a gain of 8% in the number of teachers who favor differential pay and prestige. This finding differs considerably from that of the 1984 Gallup Poll of Teachers' Attitudes Toward the Public Schools, which contained a question that was more directly worded: "How do you, yourself, feel about the idea of merit pay for teachers? In general, do you favor or oppose it?" To the directly word-

"Sometimes I think you can get too much *individual attention!"*

The teachers of today are older. They have spent more years in the classroom, and they have completed more college coursework.

ed question, 64% of the responding teachers said that they opposed merit pay.[6] Apparently, teachers' feelings about this issue are more negative when the question includes the phrase "merit pay" and the focus is solely on money.

Teachers' perceptions of their principals. Teachers' perceptions of the behaviors of their principals changed somewhat between 1964 and 1984 (Table 18). One percent fewer of the teachers in the 1984 sample felt that they could count on their principals to support them in relationships with parents. About the same proportion of teachers in both years felt that they could count on their principals to appreciate their best efforts. But 3% fewer of the teachers in 1984 agreed that their principals provide direction for the instructional program. Meanwhile, 4% more of the teachers in the 1984 sample said that their principals are effective executives, 7% more of the teachers in the 1984 sample said that their principals give them the autonomy they need to do a good job, and 12% more of the teachers in the 1984 sample said that their principals usually follow through on disciplinary problems.

THE PERSPECTIVE OF 20 YEARS

Between 1964 and 1984 the teaching force in Dade County remained the same in some ways and changed – sometimes appreciably – in other ways. The teachers of today are older. They have spent more years in the classroom, and they have completed more college-level coursework. They are slightly less satisfied with their jobs and with their schools as workplaces than teachers were 20 years ago. However, the vast majority of today's teachers express some level of satisfaction with their jobs and workplaces – despite the changes that have taken place in U.S. education since 1964 and the negative publicity that their profession has received in recent years.

In terms of their orientation toward their work, today's teachers place greater emphasis on subject matter than teachers did 20 years ago. They are more likely today to see the acquisition and use of basic skills and the development of moral behavior as important school goals, but they are less likely now than in 1964 to see imaginative problem solving and social adjustment as primary goals for the schools. At the same time as teachers have moved toward covering subject matter and helping students acquire basic skills, several scholarly studies of schooling and reports of national commissions have advocated more emphasis on critical thinking and problem solving.[7] This raises two interesting questions: What forces brought about the shift in teachers' orientation toward their work between 1964 and 1984? And what forces will be required to reverse this trend and cause teachers to focus more sharply on problem-solving and higher-order thinking skills, as some critics now advocate?

Dade County teachers also changed between 1964 and 1984 with respect to the kinds of students they prefer to teach. Today's teachers prefer students who are "nice kids," hard-working and respectful. They are less inclined than teachers were in 1964 to want to work with creative and intellectually demanding youngsters who require special efforts. If, as the recent reports on U.S. education suggest, we need to move toward academic excellence for everyone – and if, as our data suggest, teachers increasingly prefer to teach youngsters who are easy to handle, rather than those who require special efforts if they are to reach their academic limits – how are we going to reconcile this discrepancy?

Meanwhile, the recent literature on effective schools has stressed the importance of the principal's role as instructional leader. Yet our data indicate that teachers receive less instructional help from their principals and assistant principals today than they did 20 years ago. Can this trend be explained solely by the increase in resource teachers, content-area specialists, and special education consultants? Or have other forces acted on teachers and principals to shift their interactions in directions different from those advocated by many researchers who have studied school effectiveness?

If teachers today receive less instructional help from their principals and assistant principals, they receive better follow-through on disciplinary problems. The teachers in 1984 were also somewhat more likely than their counterparts in 1964 to see their principals as effective executives. This raises other questions: Can the mass of principals become effective instructional leaders while simultaneously functioning as effective executives and disciplinarians? What forces inhibit or facilitate such a

TABLE 18.
Dade County Teachers' Perceptions of Their Principals

	1964 %	1984 %
I can usually count on the principal to support me in my relationships with parents.	87.4	85.5
I can usually count on my principal to appreciate my best efforts.	78.3	77.6
My principal provides a sense of direction in the instruction given in the school.	70.7	67.6
You could call my principal an effective executive.	65.2	68.8
I can usually count on my principal to provide me with the freedom to do a good job.	80.7	87.8
The principal usually follows through on discipline problems that I bring to his or her attention.	56.9	69.2

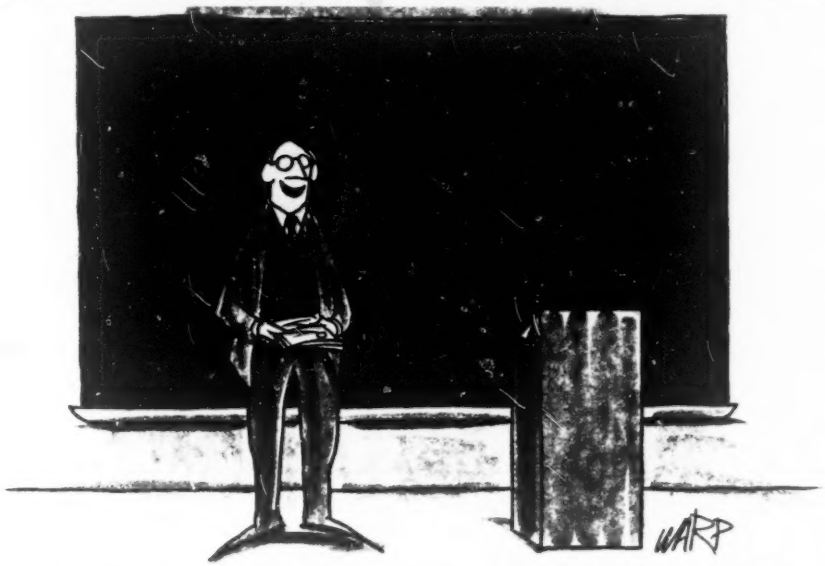
change? Why does the public view lack of discipline as the most important problem facing the schools, while teachers rank it only fourth in importance[8] and say that principals give them more support on disciplinary matters now than 20 years ago?

Meanwhile, more teachers today than in 1964 say that it is relatively easy to know when one is teaching effectively. However, in making such judgments, today's teachers rely increasingly on external sources rather than on self-assessment. One of the distinguishing characteristics of a professional is reliance on knowledge and skill (rather than compliance with external sources) in making decisions. This fact, coupled with the two findings we have just mentioned, raises several questions: Are today's teachers viewing themselves less as professionals than their counterparts did in earlier times? Are these findings indicators of the effects of "hyper-rationalization"[9] and technocratic ideology[10] on teachers? Will teachers' movement away from reliance on self-assessment ever make it possible to define academic excellence at a higher level than students' scores on tests of basic skills?

Our study of Dade County teachers in 1984 also showed that today's teachers find their work less rewarding than teachers found it 20 years ago. Today's teachers are less likely than teachers in 1964 to derive satisfaction from the respect of others or from the opportunity to wield influence. Their sense of job security and their feeling that the occupation of teaching is especially appropriate for them have both declined. But the satisfaction they derive from a schedule that allows them time away from work to fulfill other life needs has increased.

Both in 1964 and in 1984, the vast majority of teachers said that the most important rewards are intrinsic or psychic – especially a sense of "reaching" students. But 6% fewer teachers in 1984 said that intrinsic rewards are the most important kind, and 6% more teachers in 1984 said that they give priority to ancillary rewards. With regard to each of the three kinds of rewards – intrinsic, ancillary, and extrinsic – a larger number of teachers in 1984 than in 1964 said that they "receive no satisfaction from these."

How can teaching be made more rewarding? Will the currently popular "solutions" of merit pay or higher base salaries change teachers' feelings about their lack of influence and respect from the public? Are the lessened attractiveness of intrinsic rewards and the increased attractiveness of ancillary rewards (especially schedules permitting time away from work) related primarily to the demographic shift toward older teachers with greater family responsibilities? Or are these changes indicators of some more general change in teachers' attitudes toward their work? Is it possible to restore the status of teachers to previous levels or to raise their status even higher? Have the national commissions and task forces, state legislators, and the general public looked carefully enough at the complex issue of rewards, or have they focused too narrowly on salary and merit pay?



"Before I hand back your tests, I'd just like to tell you how nice it is to know that, in today's atmosphere of high-pressured competition, there are still students who don't give a damn about grades."

Our study of Dade County teachers in 1984 showed that today's teachers find their work less rewarding than teachers found it 20 years ago.

Clearly, the data from Lortie's 1964 study and from our 1984 replication of that study make possible an assessment of stability and change in teachers' views and attitudes over time. This article, the first of several reports, provides only a descriptive analysis of the changes that have occurred. But our findings raise important questions that those who seek to reform public education should ponder carefully. Through analyses of interview data and a study of forces that have affected teachers in the past, we are now seeking explanations for the changes reported here.

1. Dan C. Lortie, *Schoolteacher: A Sociological Study* (Chicago: University of Chicago Press, 1975).
2. Ibid., pp. 246-47.
3. Ibid., p. 246.
4. Louis Harris and Associates, *The Metropolitan Life Survey of the American Teacher* (New York: Metropolitan Life Insurance Co., 1984); and C. Emily Feistritzer, *The Condition of Teaching* (Princeton, N.J.: Carnegie Foundation for the Advancement of Teaching, 1983).
5. Lortie, p. 101.
6. Alec Gallup, "The Gallup Poll of Teachers' Attitudes Toward the Public Schools," *Phi Delta Kappan*, October 1984, p. 103.
7. See, for example, Ernest L. Boyer, *High School: A Report on Secondary Education in America* (New York: Harper & Row, 1983); John I. Goodlad, *A Place Called School: Prospects for the Future* (New York: McGraw-Hill, 1983); National Commission on Excellence in Education, *A Nation at Risk: The Imperative for Educational Reform* (Washington, D.C.: U.S. Government Printing Office, 1983); and Theodore Sizer, *Horace's Compromise: The Dilemma of the American High School* (Boston: Houghton Mifflin, 1984).
8. Gallup, p. 104.
9. Arthur E. Wise, *Legislated Learning: The Bureaucratization of the American Classroom* (Berkeley: University of California Press, 1979).
10. Arthur G. Wirth, *Productive Work – In Industry and Schools* (Lanham, Md.: University Press of America, 1983).

tricts, but they should be given more authority and allowed additional resources to stimulate educational improvement. While state and local administrators should not have to generate mountains of paperwork to document their programs, they should keep records adequate to justify their decisions and make those records available to anyone who is interested.

With an overhaul designed to strengthen its central purpose of improvement and to build in more accountability at all levels, Chapter 2 has the potential to become a model federal, state, and local partnership. Like most federal programs, we have it for better or worse. Let's make it better.

1. Anne Henderson, *No Strings Attached: An Interim Report on the New Education Block Grant* (Columbia, Md.: National Committee for Citizens in Education, 1983).
2. Ibid., p. 15.
3. Anne Henderson, *Anything Goes: A Summary Report on Chapter 2* (Columbia, Md.: National Committee for Citizens in Education, 1985).
4. *Nonregulatory Guidance for State and Local Education Agencies Implementing Chapter 2, ECIA* (Washington, D.C.: U.S. Department of Education, 1983).
5. Henderson, *Anything Goes. . . .*, p. 21.
6. Committee on Government Operations, *Block Grants Have Weakened Federal Programs for the Disadvantaged* (Washington, D.C.: U.S. House of Representatives, Report No. 98-581, November 1983), p. 18.
7. Council of the Great City Schools, *Federal Education Budget Policy 1980-1984: Effects on the Great City Schools* (Washington, D.C.: Council of the Great City Schools, October 1984); idem, *Trends in Federal Funding to Urban Schools: A Progress Report on the Reagan Years* (Washington, D.C.: Council of the Great City Schools, 1983); American Association of School Administrators, *Private School Participation in Chapter 2* (Roslyn, Va.: AASA, April 1984); and U.S. Department of Education Planning and Evaluation Service, "Setting the Record Straight: What ED Knows About Private School Participation Under Chapter 2," Washington, D.C., 1984.
8. *Federal Register*, vol. 47, no. 224, pp. 52370 and 52385.
9. AASA, *Private School Participation. . . .*, p. 8.
10. National Council of La Raza, "Community Access to Information on Education: A Case Study on the Education Block Grant," Washington, D.C., August 1984.
11. *A Program Evaluation of the Effectiveness of the ECIA Chapter 2* (Trenton: New Jersey Department of Education, December 1984).
12. E. H. White and Co., "Kaleidoscopes: Emerging Patterns of Response and Action in ECIA Case Studies of Chapter 2 in Selected States," Washington, D.C., 1983; and idem, "Preliminary Findings from a Study of the Implementation and Effects of Chapter 2 of the ECIA in Nine States," Washington, D.C., 1985.
13. *Lessons Learned from Past Block Grants: Implications for Congressional Oversight* (Washington, D.C.: Government Accounting Office, 1982). K

Writing for Publication: Playing to Win

by Kenneth T. Henson

How can you increase the odds that your manuscript will be accepted for publication? Heed what Mr. Henson has to say.

IN WRITING for publication there are three nevers. Never write until you have something to say. Never submit a manuscript to a journal that you haven't read. And never send a manuscript before it is ready. Though they seem painfully obvious when stated so starkly, my recent survey of 54 prominent journals in education found these to be the most common mistakes made by contributors.

There are many good reasons to write for publication. One is to communicate. An author should always provide information that a reader needs and does not have available in the existing literature. This requires knowing both the audience and the subject. Once you know what to offer an audience, write and rewrite until you communicate your message as clearly and succinctly as possible, without jargon or errors.

That rather obvious advice applies to all professional writing, whatever the field. What follows are some other suggestions that come from a 1985 survey of national journals in education. The return rate was 54 of 72 (75%). I asked the editors of these journals the questions most commonly asked by authors. The remainder of this article gives an overview of their responses, and Table 1 lists them in detail.

Should I alert the editor that I am sending the manuscript? Yes. The editor knows "what is right" for his or her journal. If the topic is wrong for the audience, or if a similar manuscript has recently been accepted, the editor will usually say so and thereby avoid a long delay that could cause the manuscript to become dated. You may either write or phone the editor. My survey found that most editors do not wish to receive query letters but that two-thirds do welcome phone calls.

How should I decide where to send the manuscript? In addition to identifying the correct audience, you should consider such factors as the journal's acceptance/rejection ratio, whether it is refereed, the size of its circulation, its typical contributors, and the length of time it takes the editor to respond to letters and manuscript submissions.

The acceptance/rejection ratio for the journals in my survey ranged from zero (some journals publish only solicited

KENNETH T. HENSON (University of Alabama Chapter) heads the Department of Curriculum and Instruction at the University of Alabama, Tuscaloosa, where he teaches writing for professional publication.

TABLE 1.

Characteristics of a Selected Sample of Education Journals

	Number of Readers	Acceptance Rate (%)	Articles Refereed (%)	Theme Issues per Year (%)	Contributors Who Are University Personnel (%)	Preferred Length (in Ms. Pages)	Min./Max. Pages	Welcome Phone Inquiries	Prefer Query Letters	Issues per Year	Articles per Year	Days to Acknowledge Receipt of Ms. (Avg.)	Days to Answer Queries (Avg.)	Weeks Required for Decision	Months Required for Publication	Effect of Photos on Acceptance	Number of Photocopies Preferred
Academic Therapy	6,000	40	No	0	20	75	8	8/10	No	Yes	5	75	2	2	1	7	None	1
Action in Teacher Education	4,000	15	Yes	90	75	80	10	1/20	No	Yes	4	14	5	5	6	4	Likely	2
Alabama ASCD Journal	1,250	66	Yes	90	-	80	15	1/15	Yes	Yes	2	16	14	14	6	3	Possibly	3
American Middle School Education	500	25	Yes	95	35	80	12	4/10	No	Yes	4	7	5	5	5	9	None	2
American Biology Teacher	7,000	33	Yes	100	10	50	8	1/18	Yes	Yes	8	120	7	4	6	9	Possibly	2
American Secondary Education	3,000	40	Yes	95	0	75	8	6/12	Yes	Yes	4	48	10	7	10	5	Likely	2
Capstone Journal of Education	400	25	Yes	95	0	80	17	5/30	No	Yes	2	12	5	3	16	6	None	3
Child Development	8,000	20	Yes	99.9	10	90	30	1/60	No	Yes	6	125	2	2	8	8	Possibly	3
Clearing House	4,100	48	Yes	100	0	75	8	2/10	No	Yes	9	112	6	20	6	7	None	2
Comparative Education Review	3,500	20	Yes	100	0	50	25	20/40	No	No	4	35	10	5	9	8	None	3
College Student Journal	1,000	35	Yes	95	0	98	5	5/10	No	Yes	4	85	30	45	6	10	Possibly	1
Computing Teacher	12,000	20	Yes	90	0	75	2+	3/10	Yes	Yes	9	54	14	10	14	3	Possibly	1
Contemporary Education	1,500	40	Yes	10	50	75	12	10/15	Yes	Yes	4	12	2	2	6	12	None	2
Core Teacher	250	75	No	0	0	75	5	--	--	--	4	12	7	7	2	6	Possibly	0
Creative Child & Adult Quarterly	--	50	Yes	--	0	95	11	1/20	No	No	4	40	7	7	4	5	Possibly	3
Delta Pi Epsilon Journal	10,000	33	Yes	100	0	95	10	--	No	Yes	2	6	2	2	13	5	None	-
Early Years/K-8	75,000	2	No	0	20	Var.	5	--	No	No	8	11	--	4	2	Var.	None	0
Eastern Education Journal	1,000	50	Yes	100	30	30	32	5/10	No	Yes	3	24	30	14	12	3	Likely	3
Education	3,500	25	Yes	90	0	95	5	5/10	No	Yes	4	85	30	45	5	10	Possibly	1
Educational Forum	50,000	20	Yes	100	0	90	18	10/20	No	No	4	35	7	7	8	9	None	3
Educational Horizons	16,500	20	Yes	10	100	80	12	1/20	Yes	Yes	4	50	4	1	16	6	None	1
Educational Leadership	80,000	8	Yes	50	100	60	10	3/20	No	No	8	130	5	2	3	--	Possibly	1
Educational Perspectives	500	10	Yes	100	100	85	10	10/12	Yes	Yes	4	20	15	15	2	4	None	0
Educational Record	10,000	15	Yes	90	75	75	12	8/18	Yes	Yes	4	48	5	2	8	6	None	1
Educational Technology	5,000	40	Yes	--	Var.	50	14	12/15	--	Yes	12	95	5	5	10	6	Possibly	1
Florida Vocational Journal	18,000	--	No	0	100	25	11	9/12	Yes	Yes	6	30	15	7	4	4	Likely	1
High School Journal	2,300	25	Yes	100	25	90	14	5/25	Yes	Yes	4	48	15	15	9	6	None	2
Impact on Instructional Improvement	475	--	Yes	--	100	33	7	6/8	Yes	No	4	38	25	25	--	--	Possibly	3
Journal of Educational Research	3,400	25	Yes	100	0	90	--	--	No	Yes	6	10	5	10	12	5	None	2
Journal of Experiential Education	--	25	Yes	100	0	95	64	1/35	Yes	Yes	4	10	14	14	14	12	None	2
Journal of Instructional Psychology	220	80	Yes	80	0	80	12	2/15	No	Yes	4	8	10	20	6	6	None	-
Journal of Reading	20,000	25	Yes	99	12	85	15	1/25	No	No	8	90	3	3	6	6	Possibly	3
Journal of Research in Science Teaching	2,000	50	Yes	100	0	95	18	--	Yes	Yes	9	90	5	0	14	10	None	2
Journal of Teacher Education	7,000	7	Yes	90	25	99	19	8/20	Yes	Yes	6	70	2	1	6	5	None	3
Middle School Journal	3,800	50	Yes	90	0	50	8	--	No	No	4	35	10	5	18	5	Possibly	5
NASSP Bulletin	3,500	20	Yes	--	100	60	11	7/15	Yes	No	9	200+	3	2	4	10	None	1
North Carolina English Teacher	900	70	No	0	12	50	8	3/15	Yes	Yes	4	40	20	10	50	10	None	-
Phi Delta Kappan	145,000	4	No	--	0	--	14	12/16	Yes	Yes	10	Var.	8	2	8	6	Possibly	0
Planning & Changing	650	40	Yes	100	--	75	5	4/6	No	Yes	4	26	7	7	5	5	None	1
Principal	23,000	10	No	0	0	15	10	5/15	Yes	No	5	75	10	21	--	2	Likely	1
Perceptual & Motor Skills	3,500	--	Yes	100	0	--	--	--	Yes	--	6	360	1	1	4	5	Possibly	4
Psychological Report	3,000	--	Yes	100	0	--	--	--	Yes	--	6	360	1	1	4	5	Possibly	7
Reading Horizons	1,000	50	Yes	90	0	85	9	2/14	No	--	4	36	7	6	7	10	Possibly	0
Reading Research Quarterly	10,000	8	Yes	100	0	85	25	--	No	No	4	32	10	7	12	12	None	4
Reading Teacher	40,000	20	Yes	99	12	85	12	1/20	No	No	8	90	3	3	6	6	Possibly	3
School Science & Mathematics	5,500	28	Yes	100	0	99	12	6/18	No	Yes	8	56	2	2	7	9	None	4
Science Education	2,000	42	Yes	99	0	--	20	4/35	No	No	5	60	30	5	14	6	Possibly	3
Social Studies	3,000	40	Yes	100	10	70	10	--	Yes	Yes	6	60	3	7	15	9	Possibly	1
Teachers College Record	3,000	10	Yes	90	25	85	25	20/30	No	No	4	24	10	7	7	7	None	2
Teacher Educator	2,000	20	Yes	100	0	2	32	5/20	No	Yes	4	22	2	2	5	6	None	0
Tech Trends	10,000	50	No	0	0	90	10	6/15	No	Yes	8	100	7	2	3	4	Likely	0
Theory Into Practice	4,000	--	Yes	100	100	75	15	10/18	No	--	4	48	10	1	8	12	None	2-3
Vitae Scholasticae	200	25	Yes	100	0	80	27	5/50	No	Yes	2	35	8	10	6	12	Possibly	2
Vocational Education Journal	50,000	20	No	0	100	70	7	6/8	No	Yes	8	120	50	10	16	Var.	Likely	1

materials) to 80%. The average acceptance rate is about 30%. More than four-fifths of these journals are refereed: 14% internationally, 33% nationally, and 21% by co-editors. Still others use a combination of these methods.

The refereeing system for U.S. journals is anything but standard, however. Because many of these journals are published on university campuses, many editors send manuscripts to professors on the same campus to be reviewed for accuracy and to be evaluated for publication. This meets the definition of refereeing espoused by some of these editors. Simply having manuscripts rated independently by the journal's staff editors appears to qualify other journals to claim refereed status. Still others send manuscripts across the country for independent rating. Even these journals, however, frequently accept manuscripts without sending them to readers for review.

The circulation size for the journals in my survey ranges from around 250 to 145,000. Most of the journals fall between 5,000 and 10,000. Seventy percent of the contributors to the journals I surveyed are university personnel. For some journals, university personnel account for only 2% to 3% of the contributors, while for others 99% of the contributors are university personnel. It takes the editors from two days to 30 days to answer a query letter. The most common time is about two weeks.

The editors in my survey acknowledge receipt of manuscripts in from one to 45 days. On the average, they acknowledge manuscripts in about a week (8.4 days). But a publishing decision takes considerably longer – anywhere from two weeks to 50 weeks. The average is about two months (8.4 weeks). Once accepted, the articles appear in most of these journals within the following 12 months; the average time from acceptance to publication is about seven months. For a few journals the lead time may exceed three years.

How long should I make my manuscript? On the average, these editors prefer manuscripts of about 10 pages, but, because the preference from journal to journal ranges from one page to 32 pages, it would be worthwhile to determine the preferred length for each particular submission. Be aware that the preferred length is strongly tied to economics. For example, one printed page in the *Phi Delta Kappan* costs upwards of $650. Thus the difference in the cost of printing a six-page article and a one-page article exceeds $3,000. If your chosen journal is not included in Table 1, count the number of words in an article of average length in that journal and divide by 300. This will give you an approximate length in manuscript pages; do not exceed this length.

Should I write for theme issues? In recent years, more and more education journals are planning issues around central themes. Almost half of the editors in my survey (48%) reported that their journals plan one or more theme issues per year; some publish *only* theme issues. To help contributors, most of these journals list upcoming themes and the deadlines for submitting articles on each theme. Using this information to time your submission or to redirect your manuscript when you learn that a journal has recently published a theme issue on your topic could help improve your acceptance rate and avoid unnecessary delays.

In addition to the original copy, how many photocopies should I send? This depends on the journal. Twenty percent of the journals in my survey want one additional copy; 23% want two; 23% want three; 4% want four; and 4% want five. One journal wants up to seven additional copies. The remaining 24% don't want any additional copies. Most journals provide this information on their inside covers. If the journal to which you wish to submit does not, you can get that information by requesting the author's guidelines from the editor or, in some cases, by consulting such reference sources as *Writer's Market*.

Would sending photos help me get my manuscript accepted? Sometimes sending photos does help. Nineteen percent of the editors I surveyed said that photos could help. Another 13.5% said that photos are likely to help. If you submit photos, however, send black-and-white glossy prints only. The preferred size depends on the individual journal. Photos that portray action in a natural setting are usually preferred. If human subjects are identifiable, you should ask them to sign permission forms authorizing you and the journal to use their pictures. Parents or guardians should sign for minors. On a piece of masking tape affixed to the back of each photo, indicate the name of the photographer who is to receive credit for the photo.

What are some additional ways that I can enhance my chances of getting my manuscript accepted? Read several recent issues of the journal to which you intend to submit a manuscript. Direct your article to the readers, not to the editor. Write simply and clearly about a topic that you like and understand. Be sure that you are contributing some new knowledge or technique or that you are approaching an issue from a distinctive perspective. Then proofread your manuscript and edit out the jargon, superfluous words, and errors. Use such resources as Table 1 to help you tailor your manuscript to the demands of your selected journal and to the needs of its readers.

With all these variables controlled, your manuscript and its fate will be in the hands of the editors. If it is rejected, remember that even editors make mistakes. Try again. K

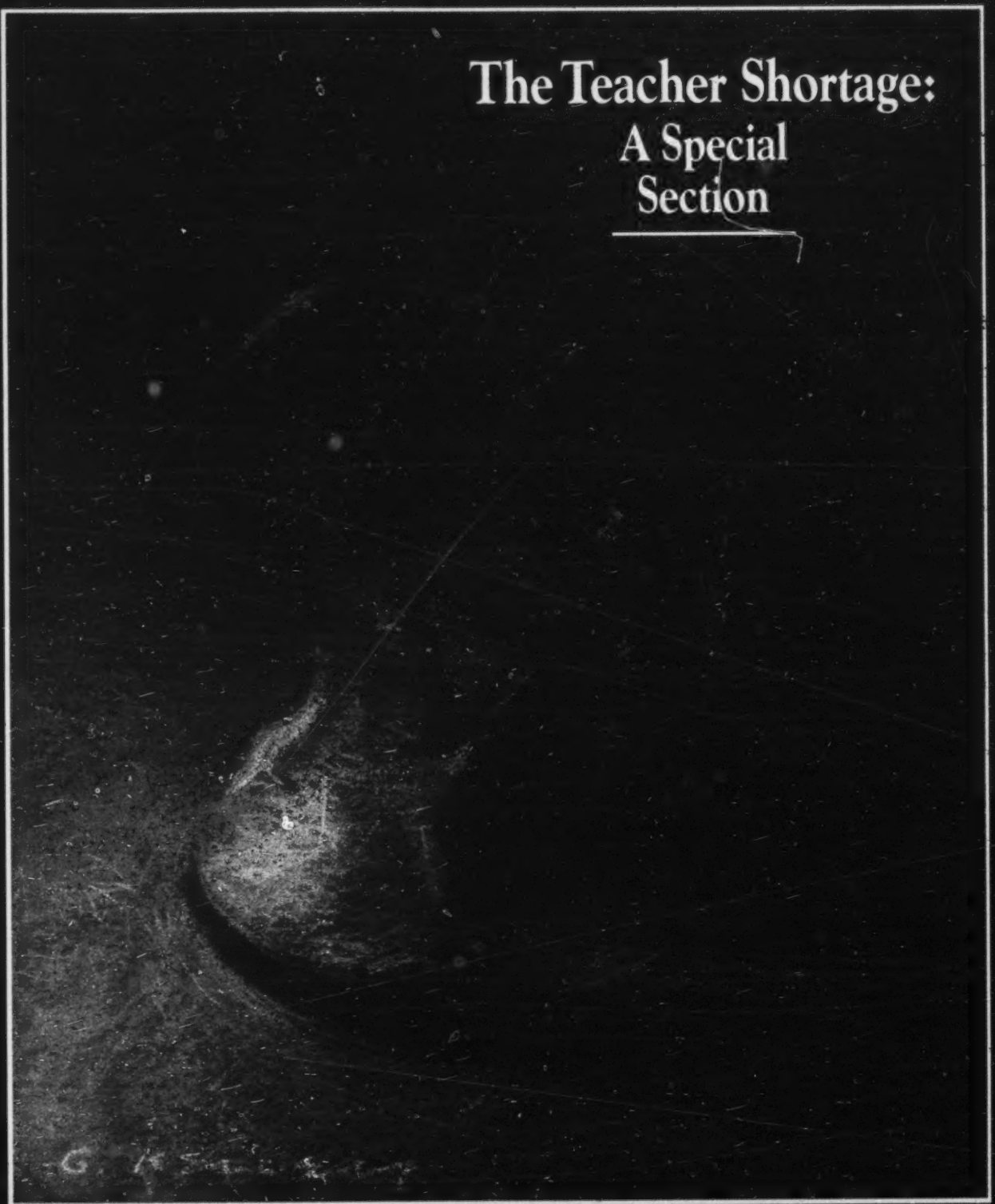
"Robby wants to show us a trick he's learned to do – can you watch, Herb? . . . Herb?"

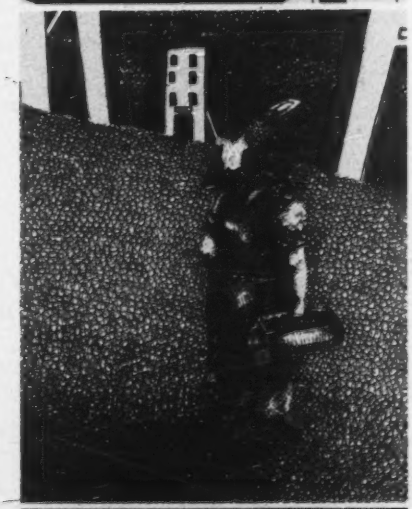
AR680-1A

PHI DELTA KAPPAN JUNE 1986

KAPPAN

The Teacher Shortage: A Special Section





Phi Delta June 1986 • Volume 67 • Number 10

KAPPAN

FEATURES

703 **Inside the Citadel in 1940,** *by Robert L. Hampel*

708 **Collaborative Links with the Community: Lessons from Exemplary Secondary Schools,** *by Bruce L. Wilson and Gretchen B. Rossman*

712 **Toward a Comprehensive Strategy for Addressing the Teacher Shortage,** *by Willis D. Hawley*

719 **Licensing Teachers: Lessons from Other Professions,** *by Martin Haberman*

723 **And Let the Air Out of the Volleyballs,** *by Gary D. Watts*

725 **Emergency Certificates, Misassignment of Teachers, and Other 'Dirty Little Secrets,'** *by Robert A. Roth*

728 **A Radical Proposal for Teacher Preparation,** *by Edwin J. Meyer*

731 **Privatization and Public Education,** *by Myron Lieberman*

735 **A Diamond in the Rough?,** *by James A. Mecklenburger*

738 **Instructional Design Flaws in Computer-Assisted Instruction,** *by Julie S. Vargas*

745 **Is CAI Cost-Effective?,** *by Henry M. Levin and Gail Meister*

750 **CAI Can Be Doubly Effective,** *by Richard P. Niemiec, Madeline C. Blackwell, and Herbert J. Walberg*

752 **Solving Educational Problems Through Partnerships,** *by Alice McDonald*

754 **No Home Working – Again!,** *by John W. Dougherty*

DEPARTMENTS

698 **THE EDITOR'S PAGE,** Desperately Seeking Susan

699 **WASHINGTON REPORT,** *by Anne C. Lewis*

701 **STATELINE,** *by Chris Pipho*

756 **PAR,** Suicide Among the Young, *by Deborah Burnett Strother*

760 **DE JURE,** Supreme Court Ducks Another Important Case, *by Thomas J. Flygare*

761 **IN EUROPE,** The Place of Europe's Women, *by Michael G. Bruce*

763 **PROTOTYPES,** The School/Community Survey: A Useful Tool in Improving Education, *by Terri V. Hunt, Marla J. Weatherl, and Deborah A. Verstegen*

764 **BOOKS,** *Laurence D. Brown*

766 **NEWSNOTES**

769 **BACKTALK**

773 **INDEX, Vol. 67**

Cover Illustration: Gary Kelley

Is CAI Cost-Effective?

Illustration by Leslie Cober

by Henry M. Levin and Gail Meister

Computer-assisted instruction is relatively cost-effective, affirm the authors, but not as cost-effective as peer tutoring. Educators should not blindly assume that CAI is tops.

IN THE 1980s, computers and education have become inextricably linked in the minds of citizens and on the agendas of school reformers. Parents worry about whether their children are "computer literate," educators worry about tracking the latest trends in technology, and students worry about gaining access to computers, both for their increasingly sophisticated games and for more serious pursuits.

Amid all this new interest in the uses of computers in education, one of the oldest uses of computers in schools, computer-assisted instruction (CAI), has come in for renewed attention. In CAI, computers are used as supplements to conventional teacher-led instruction; CAI is now used across a range of subjects in most U.S. schools.

Reduced costs and improved performance are two compelling reasons for increasing the use of computers in instruction. The price of microcomputers has fallen by 50% since 1980. At the same time, a tremendous surge in the development of educational soft-

HENRY M. LEVIN is a professor of education and director of the Institute for Research on Educational Finance and Governance, Stanford University, Stanford, Calif., where GAIL MEISTER is a research assistant.

ware has occurred. In addition to drill-and-practice programs to supplement instruction, software is now available for a wide variety of courses, including programming, logic, music, graphics, and mathematics. Despite the shrinking funding base for education in the late 1970s and early 1980s, most schools in the U.S. now have at least one microcomputer, and a great many have computer laboratories with 20 or more microcomputers.

In the rush to join the computer revolution, however, schools have not always taken the time to plan carefully. Often the extent of financial planning that precedes a decision to buy computer equipment is limited to making sure that sufficient funds are available to purchase a specified number of microcomputers, with no thought given to how they will be used. Moreover, many educators assume that the purchase of computers represents the principal cost of providing computer-assisted instruction; in fact, the purchase price is only a small portion of the cost.[1] As a result, many schools have purchased computers only to find them underused, or even unused, for want of trained personnel, good software, maintenance, or secure facilities in which to house the machines.

CAI is only one of many instructional approaches, and the cost-effectiveness of CAI is the degree to which it can provide instruction more efficiently than alternatives, as reflected in the cost of reaching a given level of educational achievement or meeting other educational goals.[2] Although many advocates claim that the drastic reductions in the costs of microcomputers have made CAI highly cost-effective, rarely are their claims supported with actual data. Our analysis here will compare figures for the costs and cost-effectiveness of CAI with those for three other interventions.

COSTS OF CAI

When parents, educators, and policy makers refer to the costs of CAI, they are often thinking only of the purchase price of the microcomputers to be used. Since microcomputers with instructional capabilities can be bought for less than $1,000 and will last for several years, CAI appears to be a very inexpensive instructional option. By this logic, a school with 1,000 students could acquire 50 computers to be used for instruction for something less than $50,000. That price would be equal to the average annual salaries of two teachers. Moreover, since the computers will last for several years, the cost in any one year would be only a fraction of the total cost. For example, if the machines were used for only three years, their apparent cost would be about $17,000 per year, or $17 per student per year. Clearly, such a price seems a very modest part of the $3,000 or so spent on each student each year.

Despite assertions that CAI is more cost-effective than other instructional interventions, little or no evidence exists to support this notion.

The problem with this kind of calculation, however, is that microcomputers by themselves are only one element in a more complex system for delivering instruction. In addition to the machines themselves, schools need a secure facility in which to house them, appropriate software, knowledgeable personnel, provisions for maintenance, and other support services. The proper way to determine the cost of CAI is first to ask what ingredients are needed to deliver a particular level of instructional services, then to determine the costs of those ingredients, and finally to determine who pays for them. This "ingredients approach" takes account of all the costs of CAI and apportions them among the sources of support.[3]

An ingredients approach entails a careful analysis of the CAI program in question, whether proposed or already in place. This analysis requires a description of the number of students, the duration of the instruction, and the necessary personnel, facilities, hardware, software, training, and maintenance. After a detailed inventory of these ingredients has been completed, their costs are determined. Then the costs can be allocated to the different constituencies who will provide the ingredients or pay for the services.

It is particularly important to allocate the costs of CAI hardware and facilities over their respective lifetimes. For example, if a renovated facility is expected to last for 20 years and the hardware for only five years, it is inappropriate to calculate the cost per student solely according to the expenses incurred in the year in which the facilities are renovated or the computers purchased. Although adequate financing is necessary for start-up expenses, the actual cost of CAI per student will depend on the number of years that the facilities and equipment are used. Fortunately, easily applied techniques are available for converting expenditures on equipment and facilities with different lifetimes into annual cost estimates.

The first systematic study of the costs of CAI was carried out in 1978 for the widely used drill-and-practice approach of the Computer Curriculum Corporation (CCC).[4] The CCC delivery system uses a computer laboratory with a minicomputer and 32 student terminals. The lab is staffed by an instructional coordinator and part-time instructional aides. (We should note here that cost reductions and performance improvements for microcomputers have paralleled those for minicomputers.) Using the ingredients approach, we calculated that a daily session of 10 minutes of drill and practice would cost about $136 a year per student. Even considering the relatively high cost of computer equipment in 1978, the hardware accounted for only about one-quarter of the total annual costs for delivering CAI services.

In 1984 the costs were reestimated, using the most recent costs for computer hardware, software, and maintenance and using 1980 costs for personnel, facilities, and other ingredients.[5] The 1980 data were part of a standard database on costs used at the time for comparing different types of interventions. However, the cost information for the hardware and software was updated to 1984 to take account of the tremendous reductions in these costs between 1980 and 1984.

We found that the cost of providing the same CAI services as in the earlier study had fallen to about $120 per student per year. The relative cost of the computer hardware had decreased drastically between 1978 and 1984, from one-quarter of the total annual cost of the CAI services to one-ninth. This finding highlights the importance to school decision makers of considering costs other than hardware when estimating the cost of a proposed CAI program. Almost 90% of the cost of CAI is associated with personnel, software, training, and other factors. In 1984 personnel costs alone accounted for about

38% of the total, and curriculum and other software costs accounted for another 21%.

These figures suggest the limits to future reductions in the cost of CAI. Reductions in hardware, however large, can affect only about 11% of the total cost of CAI.

Let us try a hypothetical exercise in arithmetic. Assume that hardware costs in the future will decline so drastically that computers will be given away at no cost whatever. Although this is admittedly an absurd assumption, it helps illustrate the limit to which reduction in hardware costs would reduce the overall cost of CAI. The use of free computers would reduce the cost of a CAI program by 11% at most. However, since the costs of other ingredients would tend to rise, the total reduction in costs would be somewhat less. Thus claims that drastic reductions in the cost of microcomputers will have an equally drastic impact on the cost of CAI are vastly overstated.

An additional finding of our research is that microcomputers linked in a local area network do not appear to be cheaper than a central minicomputer connected to data-entry terminals. Microcomputer networks are designed to take advantage of capabilities for sharing such peripheral equipment as storage devices and printers, for communicating among users, for monitoring instruction and obtaining progress reports for groups of students, and for storing student records. Although minicomputers have been performing these tasks for some time, many educators have assumed that inexpensive microcomputers linked with a local network would perform the same tasks for less. We conclude, from our study of currently available technologies, that minicomputers with terminals are actually less expensive and more reliable than comparable configurations using microcomputers.[6]

FOUR INTERVENTIONS

Despite vigorous assertions that CAI is more cost-effective than other instructional interventions, little or no evidence exists to support this notion. Our research, conducted with Gene Glass under the auspices of the Institute for Research on Educational Finance and Governance at Stanford University, compared the cost-effectiveness of four educational interventions designed to raise mathematics and reading achievement at the elementary school level.[7] The interventions we studied were CAI, cross-age tutoring, reduced class size, and a longer school day. Each of these interventions is an attempt to improve educational achievement, and CAI and longer school days were prominent among the recommendations of the recent national reports on reforming education.

Cost-effectiveness analysis requires detailed and precise data on both costs and effects. Effectiveness data can be derived only from a systematic evaluation of a particular intervention in a particular setting. Although there have been a great many approaches to CAI, few have been evaluated. The single most prominent use of CAI has been to provide drill and practice to reinforce classroom instruction in mathematics and reading. Educators have had extensive experience with this use of CAI, and competent evaluations are available.

The relative cost-effectiveness of CAI could improve as new curriculum materials and new applications continue to be developed.

CAI. We chose the drill-and-practice curricula of the Computer Curriculum Corporation as the specific CAI approach for our cost-effectiveness evaluation. The Educational Testing Service evaluated the effectiveness of this approach in a four-year experiment in the Los Angeles Unified School District.[8] To derive cost estimates, we identified each program ingredient and determined its cost. As we said above, we used 1984 prices for the computers, software, and maintenance, but we used 1980 prices for the personnel, facilities, and other factors.

Our cost estimation probably understates slightly the cost of CAI by merging the latest hardware costs with relatively low personnel costs. Unlike CAI, the cost of the three remaining educational interventions did not experience dramatic changes over the period, and so their cost estimates are based uniformly on 1980 prices. We settled on the 1980 figures in order to provide a consistent set of national costs on which to base our analysis across interventions.

Cross-age tutoring. Cross-age tutoring refers to the use of adults or older students as tutors. This practice is attractive because considerable evidence suggests that, when students are used as tutors, both tutor and tutee gain in academic achievement from the tutoring relationship.

We chose for our comparison the Cross-Age Structured Tutoring Program for Reading and Mathematics, used in the public schools of Boise, Idaho.[9] Staffing for the program in each elementary school of about 300 to 400 students included adults as tutor/managers in reading and mathematics and as one-on-one tutors. The adults worked with the children in the upper

"We will now begin our computer training workshop for teachers."

grades, while children in the upper grades worked with second- and third-graders. The services of a district coordinator were allocated to the school on a part-time basis. Daily tutoring sessions used commercially available software and lasted about 20 minutes. Each school had about 60 students tutoring 60 younger children and about 26 older students assigned to adult tutors.

Increasing instructional time. Increasing instructional time was a centerpiece of the recent national calls for reforming education. The specific intervention we chose to evaluate was the addition of one hour of instruction per day, half of this time devoted to reading and half of it to mathematics. Data on the effectiveness of this increase in instructional time were derived from the most important evaluation of the effects of time on learning: the Beginning Teacher Evaluation Study (BTES), which was sponsored by the National Institute of Education.[10]

Reducing class size. Reducing class size has traditionally been seen as a means of improving instruction. We derived our evidence regarding the effects of class size on achievement in reading and mathematics from a meta-analysis of nearly 80 evaluations, conducted by Gene Glass and his colleagues.[11] Specific estimates were made for increases in mathematics and reading achievement associated with reductions in class size from 35 to 30 students, from 30 to 25 students, from 25 to 20 students, and from 35 to 20 students. We estimated costs for the additional teachers, classrooms, and classroom resources required for this intervention.

COST-EFFECTIVENESS

Effectiveness. Table 1 shows the estimated effectiveness of each of the four interventions in months of additional student gain in achievement per year of instruction. That is, each of the interventions represents a supplement to existing instruction, so the effects are evaluated in terms of the additional achievement expected for children who receive the instructional intervention relative to those who do not.

The study from which these findings are taken measured effect size in standard deviation units. However, a standard deviation of achievement at the elementary level is approximately equal to an academic year of achievement, where an academic year equals 10 months. For this reason, Table 1 presents the effectiveness results in terms of months of additional student gain in each subject; each month of gain is about one-tenth of a school year of achievement.

The CAI intervention produced a healthy result, with more than a month of student gain in mathematics and more than two months (almost a quarter of a year) in reading, for a 10-minute daily

TABLE 1.
Estimated Effectiveness of Four Educational Interventions in Months of Additional Achievement Gain per Year of Instruction

		Mathematics	Reading
CAI		1.2	2.3
Cross-age tutoring			
Peer component		9.7	4.8
Adult component		6.7	3.8
Increasing instructional time		0.3	0.7
Reducing class size			
From	To		
35	30	0.6	0.3
30	25	0.7	0.4
25	20	0.9	0.5
35	20	2.2	1.1

TABLE 2.
Annual Cost of Four Educational Interventions

		Cost in Dollars per Student per Subject
CAI		119
Cross-age tutoring		
Peer component		212
Adult component		827
Increasing instructional time		61
Reducing class size		
From	To	
35	30	45
30	25	63
25	20	94
35	20	201

TABLE 3.
Estimated Effectiveness of Four Educational Interventions in Months of Additional Achievement Gain per Year of Instruction for Each $100 per Student

		Mathematics	Reading
CAI		1.0	1.9
Cross-age tutoring			
Peer component		4.6	2.2
Adult component		0.8	0.5
Increasing instructional time		0.5	1.2
Reducing class size			
From	To		
35	30	1.4	0.7
30	25	1.2	0.6
25	20	1.0	0.5
35	20	1.1	0.6

session in each subject. Even larger effects were found for peer tutoring, however, with a gain of almost one full year in mathematics achievement and almost half a year in reading achievement. By contrast, reducing class size showed less than a month of gain in both reading and mathematics for each five-student decrement. A direct reduction from 35 to 20 students, however, was associated with gains similar to those for CAI, but with subject areas reversed. Finally, the addition of a half hour per day of instruction in each subject yielded very small gains.

Cost. Table 2 shows the estimated cost per student of each intervention, using the ingredients method of estimating costs. The cost per student per subject represents the total value of resources required to replicate each intervention for either reading or mathematics, divided by the number of students. The most costly of the interventions is adult tutoring, followed by peer tutoring and a reduction in class size from 35 to 20. The cost of CAI is about half that of peer tutoring. Reducing class size in five-student decrements or increasing instructional time by a half hour per day in each subject are the least costly interventions.

When the costs for each intervention in Table 2 are combined with the effectiveness results from Table 1, a cost-effectiveness ratio can be derived. The combination allows us to calculate the expected gains in student achievement associated with a given cost. Table 3 shows the expected gains in student achievement of each intervention for each $100 invested per pupil.

CAI is estimated to produce a gain of about one month in mathematics and two months in reading for each $100 spent per student. Peer tutoring, by contrast, is associated with almost half a year of achievement gain in mathematics and almost a quarter year of achievement gain in reading for each $100 invested per student. The other interventions show lower cost-effectiveness than peer tutoring or CAI. Indeed, though adult tutoring shows one of the largest effects, its high cost creates a cost-effectiveness ratio that is among the lowest of the four interventions.

From these results, we conclude that the specific CAI intervention we evaluated was more cost-effective than adult tutoring, reducing class size, or increasing instructional time. However, it was considerably less cost-effective than peer tutoring in mathematics and slightly less cost-effective than peer tutoring in reading. This suggests that, though CAI is a relatively cost-effective intervention, it is not necessarily the most cost-effective approach to improving mathematics and reading achievement in the elementary grades.

The relative cost-effectiveness of CAI could improve as new curriculum materials and new applications are developed. Work is under way at the Institute for Research on Educational Finance and Governance to explore the cost-effectiveness implications of rapid changes in hardware, software, instructional configurations, and approaches to implementation.

So far, however, the evidence suggests that educators should not assume blindly that CAI is a more cost-effective intervention than the alternatives. Clearly, the criteria for evaluating proposed or existing CAI and other educational interventions must depend on more than just cost-effectiveness. The evaluation criteria also include the instructional goals of a school, the resources available for reaching those goals, the proficiency of the staff in using a given intervention, and a great many other factors.

"Then one day I realized that, when the office computers are down, I'm down."

1. See, for example, Stanley Pogrow, *Education in the Computer Age* (Beverly Hills, Calif.: Sage, 1983), pp. 80-81, 215; and Israel Pressman and Bruce Rosenbloom, "CAI System Costs: Present and Future," *T.H.E. Journal*, March 1984, pp. 94-98.

2. Henry M. Levin, "Cost-Effectiveness Analysis in Evaluation Research," in M. Guttentag and E. Struening, eds., *Handbook of Evaluation Research, Vol. 2* (Beverly Hills, Calif.: Sage, 1975), pp. 89-122.

3. This method and rationale are described in Henry M. Levin, *Cost-Effectiveness: A Primer* (Beverly Hills, Calif.: Sage, 1983).

4. Henry M. Levin and Louis Woo, "An Evaluation of the Costs of Computer-Assisted Instruction," *Economics of Education Review*, Winter 1981, pp. 1-26.

5. Henry M. Levin, Gene V Glass, and Gail R. Meister, *Cost-Effectiveness of Four Educational Interventions* (Stanford, Calif.: Institute for Research on Educational Finance and Governance, Project Report No. 84-A11, May 1984). This comparison is discussed in greater detail in Henry M. Levin, *Costs and Cost-Effectiveness of Computer-Assisted Instruction* (Stanford, Calif.: Institute for Research on Educational Finance and Governance, Project Report No. 84-A21), pp. 18-22.

6. Levin, Glass, and Meister, pp. 24-25.

7. Ibid.

8. Marjorie Ragosta et al., *Computer-Assisted Instruction and Compensatory Education: The ETS/LAUSD Study, Final Report* (Princeton, N.J.: Educational Testing Service, Project Report No. 19, April 1982).

9. Independent School District of Boise City, "1982-83 Evaluation Report, PROJECT INSTRUCT, Chapter 1 Program," July 1983; idem, "Submission to Joint Dissemination-Review Panel (J.D.R.P.), Cross-Age Structured Tutoring Program in Math," mimeo, n.d.; and idem, "Submission to Joint Dissemination-Review Panel (J.D.R.P.), Cross-Age Structured Tutoring Program in Reading," mimeo, n.d.

10. Charles W. Fisher et al., "Teaching Behaviors, Academic Learning Time, and Student Achievement: An Overview," in C. Denham and A. Lieberman, eds., *Time to Learn* (Washington, D.C.: National Institute of Education, 1980), pp. 7-32; and Nancy Karweit, *Time-on-Task: A Research Review* (Baltimore: Center for Social Organization of Schools, Johns Hopkins University, Report No. 332, 1983).

11. Gene V Glass and Mary L. Smith, "Meta-Analysis of Research on Class Size and Achievement," *Educational Evaluation and Policy Analysis*, vol. 1, 1979, pp. 2-16.

CAI Can Be Doubly Effective

by Richard P. Niemiec, Madeline C. Blackwell, and Herbert J. Walberg

Hold on, say the authors. Computer-assisted instruction is about twice as effective as peer tutoring. They explain why.

HENRY LEVIN and Gail Meister propose cost-effectiveness as a way of choosing among instructional alternatives. They conclude from their analysis that computer-assisted instruction (CAI) is not the most cost-effective instructional intervention and that peer tutoring yields the highest achievement gains per $100 paid for instruction. Accurate, broad-based, and up-to-date estimates of these interventions, however, indicate that CAI is about twice as cost-effective as peer tutoring.

To determine ratios of cost-effectiveness, Levin and Meister first calculated the ratio of effects of each of several instructional methods to its respective costs. The numerator of that ratio is the effect of the intervention expressed in standard deviation units or effect sizes. For ease of interpretation, they have converted these effect sizes into achievement gains in months (and assumed that one-tenth of a standard deviation equals a one-month grade equivalent – a questionable procedure but one that we retain here for purposes of this discussion). The denominator of the ratio is an estimate of the yearly per-pupil cost of the intervention. This ratio is multiplied by 100 to estimate the achievement gains per $100. The accuracy of Levin and Meister's estimates is crucial. First, consider their estimation of the costs of implementing CAI and peer-tutoring programs. Using their ingredient method, they estimate that the per-pupil cost of CAI is $119 per year and that the per-pupil cost of peer tutoring is about $212 per year. These estimates are derived from a combination of 1984 computer hardware costs and 1980 costs for all other ingredients. These estimates assume that 1980 costs adequately reflect the true costs of implementing the programs in 1986.

Although this may be a valid assumption, other estimates that make different assumptions will yield different costs. For example, increasingly better software could make teacher supervision of and assistance to students less time-consuming; teacher aides, who are paid significantly less than teachers, could take responsibility for some tasks; and the computer could keep track of individual student records. Such changes might reduce the costs of labor considerably.

The estimates of the effects of the interventions are equally critical to Levin and Meister's argument, and, as with cost estimates, changes in assumptions profoundly change the cost-effectiveness ratios. Levin and Meister have chosen two studies that they claim are representative of the effects of peer tutoring and CAI. For tutoring, they have chosen an undated, unpublished study of the Boise, Idaho, Tutoring Program. The unnamed authors of this evaluation apparently indicated that the program produced an effect size of .97 in mathematics – almost a full year's gain – and of .48 in reading – almost half a year's gain. The combined effect size is .73, or more than a seven-month gain in achievement.

But how representative is the Boise study? Would a seven-month gain in achievement be an average gain for a tutoring program? As 65 independent evaluations have shown, the average ef-

RICHARD P. NIEMIEC is director and MADELINE C. BLACKWELL is assistant director of the Computer-Assisted Instruction Laboratory at Southeastern Louisiana University, Hammond. HERBERT J. WALBERG (University of Chicago/DePaul University Chapter) is Research Professor of Education at the University of Illinois, Chicago.

TABLE 1.
Estimated Effects In Months

	Levin and Meister	Our Estimates
Peer Tutoring		
Mathematics	9.7	6.2
Reading	4.8	2.1
Combined	7.3	4.2
CAI		
Mathematics	1.2	5.4
Reading	2.3	4.2
Combined	1.8	4.8

TABLE 2.
Cost-Effectiveness Ratios

	Levin and Meister	Our Estimates
Peer Tutoring		
Mathematics	.46	.29
Reading	.23	.10
Combined	.35	.20
CAI		
Mathematics	.10	.45
Reading	.19	.35
Combined	.15	.40

Illustration by Leslie Cober

fect of tutoring programs is .40 in mathematics, or only a four-month gain.[1] However, only 15 of these 65 studies reported effects larger than .5. This suggests that the median, or typical, effect is substantially lower than .40. Fully 50 of the studies reported "small or trivial effects of tutoring."[2] This quantitative review indicated that the average effect sizes were .62 for mathematics tutoring and .21 for reading (see Table 1).

Levin and Meister's choice as a representative study of the effects of CAI is the Educational Testing Service/Los Angeles Unified School District's evaluation of the Computer Curriculum Corporation drill-and-practice program. This single report indicates that CAI increased mathematics achievement by about .12, or slightly more than one month; it increased reading achievement by about .23, or slightly more than two months. Although this study was well-documented, the authors evaluated only one form of CAI, produced by one company.

Are these effects typical of the research findings on CAI in general? A quantitative synthesis of 29 studies yielded an average CAI effect of .47, considerably higher than Levin and Meister's estimate based on a single study.[3] Our own estimate was .45, which we derived from 48 studies.[4]

It appears that Levin and Meister overestimated the achievement gains of peer tutoring and underestimated those of CAI.

Given these findings, it appears that Levin and Meister overestimated the achievement gains of peer tutoring and underestimated those of CAI. More broadly-based syntheses of the research on these interactions would yield more representative estimates of the effects of the interventions. But our estimates are derived from the findings of 110 studies, in contrast to the two used by Levin and Meister in deriving their estimate.

Table 2 shows cost-effectiveness ratios calculated using our estimates of the effects of CAI and peer tutoring and those of Levin and Meister. Their estimate of cost-effectiveness suggests that peer tutoring is more than twice as cost-effective as CAI. Ours yields the opposite conclusion. Using our figures, the cost-effectiveness ratio of CAI is .40 – twice the estimated effect of peer tutoring. Levin and Meister pose the question, Is CAI cost-effective? The answer is yes. Compared to peer tutoring, adult tutoring, increasing the length of the school day, and decreasing class size, an average CAI program produces the greatest gains per $100 of instructional expenditure. Of course, further increases in the effectiveness of CAI, reduced labor costs, and particular local conditions are likely to affect the estimates as well.

1. Peter A. Cohen, James A. Kulik, and Chen-Lin C. Kulik, "Educational Outcomes of Tutoring: A Meta-Analysis of Findings," *American Educational Research Journal*, Summer 1982, pp. 237-48.
2. Ibid., p. 243.
3. James A. Kulik, Chen-Lin C. Kulik, and Robert Bangert-Drowns, "Effectiveness of Computer-Based Education in Elementary Schools," *Computers in Human Behavior*, vol. 1, 1985, pp. 59-74.
4. Richard P. Niemiec, Gordon Samson, Thomas Weinstein, and Herbert J. Walberg, "The Effects of Computer-Based Instruction in Elementary Schools: A Quantitative Synthesis," *AEDS Journal*, in press. K